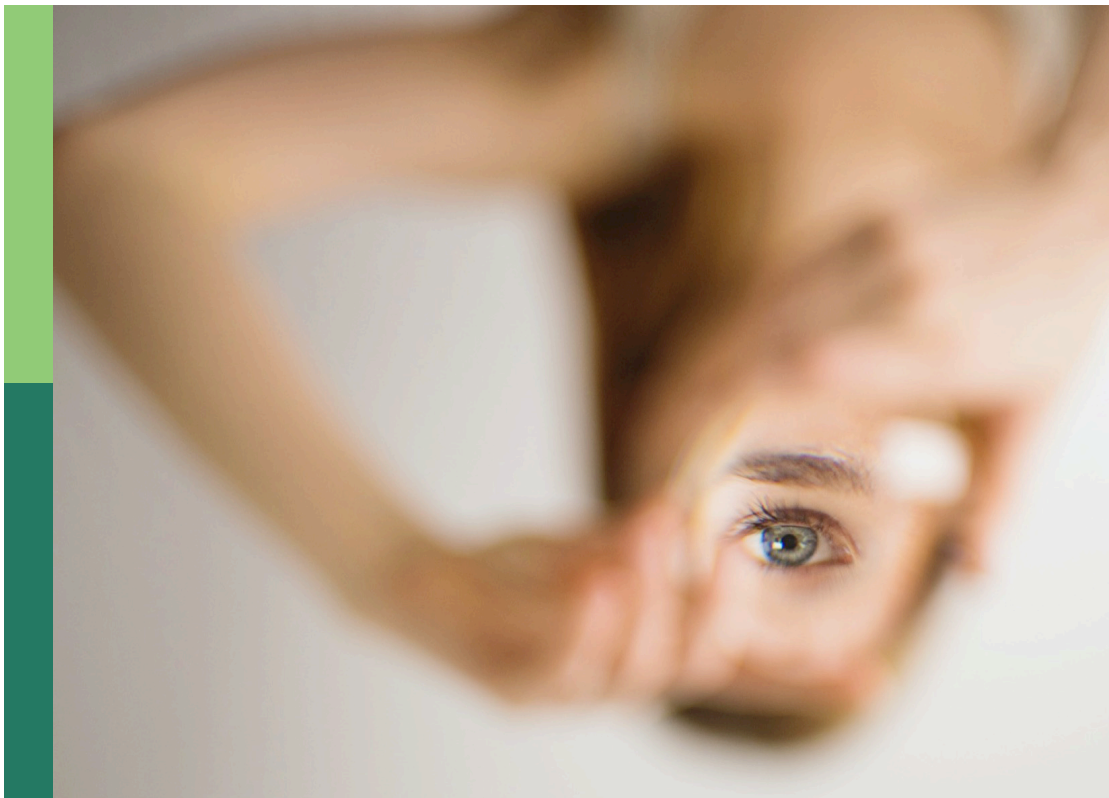
frontiers | Research Topics

Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes

Edited by
Alison Kington, Ridwan Maulana and James Ko

Published in
Frontiers in Psychology
Frontiers in Education



ISSN 1664-8714
ISBN 978-2-8325-3327-7
DOI 10.3389/978-2-8325-3327-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes

Topic editors

Alison Kington — University of Worcester, United Kingdom

Ridwan Maulana — University of Groningen, Netherlands

James Ko — The Education University of Hong Kong, Hong Kong, SAR China

Citation

Kington, A., Maulana, R., Ko, J., eds. (2023). *Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3327-7

Table of contents

frontiers | Frontiers in Education

TYPE Editorial
PUBLISHED 17 March 2023
DOI 10.3389/feduc.2023.1170854

OPEN ACCESS

EDITED AND REVIEWED BY
Douglas F. Kauffman,
Medical University of the Americas – Nevis, United States

*CORRESPONDENCE
Ridwan Maulana
✉ r.maulana@rug.nl

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

RECEIVED 21 February 2023
ACCEPTED 27 February 2023
PUBLISHED 17 March 2023

CITATION
Maulana R, Kington A and Ko J (2023) Editorial: Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes. *Front. Educ.* 8:1170854. doi: 10.3389/feduc.2023.1170854

Editorial: Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes

Ridwan Maulana[1]*, Alison Kington[2] and James Ko[3]

[1]Department of Teacher Education, University of Groningen, Groningen, Netherlands, [2]School of Education, University of Worcester, Worcester, United Kingdom, [3]Department of Education Policy and Leadership, The Education University of Hong Kong, Hong Kong, Hong Kong SAR, China

KEYWORDS

effective teaching, measurement, correlates, outcomes, characteristics

Editorial on the Research Topic
Effective teaching: Measurements, antecedents, correlates, characteristics, and links with outcomes

A growing desire for improved quality in education has been evident throughout the world. While some scholars are focused on the technical aspects of teaching, others are exploring how, why, when, and for whom certain factors contribute to educational achievement and success. Effective, quality education suggests the need to understand the key stakeholders within the education system—the teachers and students. However, although research on effective teaching has existed for over half a century, the knowledge base is limited. There has been a heavy reliance on using student achievement as an outcome criterion but, more recently, non-cognitive psychological outcomes have been acknowledged as important factors for improving the effectiveness and quality of education. This Research Topic included 14 articles that contribute to advancing the knowledge base on teaching effectiveness and its complex measurements, antecedents, correlates, characteristics, and links with cognitive and non-cognitive outcomes. We intend that the Research Topic will stimulate international discussion and inspire future work for advancing the field and its impact on research, practice, and policy.

1. Introduction

This Research Topic is grounded in the relatively recent Global Effective Teaching and Learning Network (GETLIN) network[1]. This network was established as a result of the International Comparative Analysis of Learning and Teaching (ICALT3/Differentiation) project led by the University of Groningen (2015–2022). The project brought together researchers, educators, practitioners, and policymakers from over 17 education system to share knowledge and exchange ideas about teaching quality improvement. This project paved the way for continuous cooperation in effective teaching among key educational

1 GETLIN was initially established during the final ICALT3/Differentiation conference from 26 – 28 July 2021 in Groningen, the Netherlands.

stakeholders in numerous countries across the five continents. This Research Topic is partly the result of this international collaboration, which complements two other book publications on effective teaching in an edited volume titled *Effective teaching around the world: Theoretical, empirical, methodological and practical insights* (Maulana et al., 2023a) and a monograph titled *Differentiated instruction in teaching from the international perspective: Methodological and empirical insights* (Maulana et al., 2023b).

Teaching is a complex activity in a highly complex environment. Although our understanding of effective teaching has significantly widened over the last century, major gaps remain. What constitutes "effective" teaching differs from one context, time, and group to another. As Ko et al. (2014) noted, defining effective teaching can be complex and controversial due to its associations with professional competency and system accountability. In general, the scope of effective teaching can be defined in two ways. The first definition focuses on processes *inside* the classroom and is generally referred to in terms of the focus on teacher behaviors and classroom processes that promote better student outcomes. The second definition takes a broader view, including factors *beyond* classroom processes, based on the notion that teachers' work is not restricted to instruction alone (Ko et al., 2014). In the 21st century, the role of a teacher extends beyond classroom processes, and includes, inter alia, collegial coaching and cooperation, teacher leadership, and engagement in continuous professional development.

Scholars tend to agree that effective teaching invokes *change* in outcomes. This claim supports the work of Maulana et al. (2023a). Previously, emphasis was put on change in cognitive outcomes (i.e., test scores). In contemporary education, the importance of viewing change in terms of non-cognitive outcomes has been identified. In addition, viewing change not only from student outcomes but also teacher outcomes has been highlighted. This Research Topic was chosen based on the premise that effective teaching is closely linked to teachers as crucial agents for change in students, and their professional, learning. It discusses multiple issues of effective teaching, including its measurements, antecedents, correlates, characteristics, and relationships with outcomes.

For this Research Topic, we have attempted to identify the developing field of effective teaching by looking at this complex and admittedly contested concept from a wide variety of perspectives worldwide. We hope that this Research Topic will stimulate further discussion on effective teaching and invite scholars to connect and cooperate to contribute to improving the knowledge base by bringing unique expertise, experiences, and interests from multiple contexts around the globe.

2. The papers

Table 1 provides an overview of papers included in this Research Topic (see Table 1). Although the topic focuses on effective teaching, it should be noted that the papers address both direct and indirect forms of effectiveness. We interpret this variation as a representation of effective teaching, which is complex by nature and encompasses broad concepts. As can be seen in Table 1, the papers included in this Research Topic cover diverse components of effective teaching, methods, and education levels and originate from diverse education systems. It is important to note that all the papers address multiple issues (e.g., measurements, characteristics, and links with outcomes) of effective teaching. Table 2 summarizes specific components of the 14 papers concerning effective teaching.

Based on the broader and contemporary definitions of effective teaching, the papers can be divided into groups based on two dimensions of effective practice; (i) within-classroom processes and (ii) factors beyond classroom processes.

2.1. Effective teaching within the classroom processes

A total of seven articles address effective teaching within the classroom processes. All of these uniquely present various dimensions of effective teaching with some conceptual overlaps.

The paper by Yang et al. presents a systematic literature review of the link between teachers' autonomy support and student engagement using the PRISMA method, a scoping systematic review approach. Primary, Secondary, and Higher Education studies from various education systems are represented. Their study sheds light on trends regarding autonomy support and engagement in terms of landscapes, methodology, characters, patterns of studies, and autonomy-supportive strategies. They found limitations in the current literature on autonomy support and engagement and pointed to the need for further longitudinal studies on specific autonomy-supportive teaching strategies to maximize student engagement.

Li et al. report on university students' perceptions of motivational lesson climate, engagement, and achievement when enrolling in English classes in Northeastern and Mid-eastern China. Student questionnaires were used to measure perceptions of motivational climate and engagement. The final English test score was used as a measure of achievement, and exploratory and confirmatory analysis, regression, and path analyses were employed. This study reveals the important predictive power of perceived motivational climates on perceived engagement. The paper concludes that English language teachers in China exhibit certain aspects of effective motivational climates, but other aspects of motivational climates need to be improved to maximize student engagement and achievement.

The paper by Ouwehand et al. examines the relationship between school composition (i.e., student socio-economic status (SES) and migration background), workload, and teachers' utility values toward teaching and teaching quality (i.e., classroom management, cognitive activation, clarity of instruction, and teacher-student relationship) in Dutch secondary schools using TALIS 2018 data. Data were gathered using teacher and school leader questionnaires and were analyzed using a multilevel modeling approach. The study reveals that certain dimensions of social utility value matter in predicting teaching quality. SES was also shown to relate to certain dimensions of teaching quality, suggesting that different dimensions of SES matter more for specific dimensions of teaching quality.

TABLE 1 A summary of papers included in the Research Topic.

Research Topics	Type	Approach	Location	Education level	Method and measure	Focus
1. Yang et al.	Systematic review	Scoping systematic review	International	Primary, Secondary, Higher Education	Longitudinal, PRISMA method	Teachers' autonomy support, Student engagement
2. Li et al.	Original research, 2 studies	Quantitative, factor analysis (EFA and CFA), regression analysis, path analysis	Northeastern China, Mid-eastern China	Higher education	Survey using student questionnaire	Motivational climate, Student engagement, Student achievement
3. Ouwehand et al.	Original research	Quantitative, multilevel analysis	Netherlands	Secondary schools	Survey using teacher and school leader questionnaires	Teaching quality: classroom management, cognitive activation, teacher-student relationship school population composition: SES, migration background workload, teacher utility values toward teaching
4. Gomariz et al.	Original research	Quantitative, factor analysis (EFA, CFFA), structural equation modeling	Spain	Early childhood, primary, secondary education	Survey using teacher questionnaire	Teacher facilitation of family participation: communication with the center, participation in school activities, sense of belonging, involvement in the home, involvement in the pa and the school board, community participation, training
5. Melguizo-Garín et al.	Brief research report	Quantitative, descriptive, correlational, multiple regression analyses	Spain	Higher education	Survey using student questionnaire	Analysis of cooperation in higher education: conception of group work, usefulness of group work, planning of group work by teachers, criteria for organizing groups, group norms, internal functioning of groups, effectiveness of group work, student satisfaction with PBL
6. Johansson et al.	Original research	Quantitative, multiple regression, multilevel growth modeling	Sweden	Primary, secondary, higher education	Secondary data analysis, teacher register	University credits, preservice training, in-service training
7. Pan et al.	Original research	Qualitative and quantitative, systematic literature review (inductive content analysis), expert survey (delphi method)	International	No information, most likely general	Inductive content analysis, expert survey (Delphi method)	Game-based learning, Smart classroom
8. Liu	Original research	Quantitative, fixed-effect modeling	China	No information, most likely general	Secondary data survey, Pupil learning gains	Educational credentials, Pupil learning gains

(Continued)

10.3389/feduc.2023.1170854

TABLE 1 (Continued)

Research Topics	Type	Approach	Location	Education level	Method and measure	Focus
9. Gencoglu et al.	Conceptual analysis	Qualitative, theoretical analysis	International	General	Conceptual/theoretical analysis	Student perceptions, variability, teaching quality, individual and cultural factors
10. Falcón-Linares et al.	Original research	Quantitative, correlation, qualitative, grounded theory	Spain	Secondary education	Survey using semi-structured questionnaire	School counselor, school tutoring, vocational guidance, secondary education, qualitative research
11. van de Pers and Helms-Lotenz	Original research	Quantitative, longitudinal, linear regression analysis	Netherlands	Secondary education	Questionnaire survey of teachers	Induction, support, beginning teachers, school context
12. Collie et al.	Original research	Quantitative, CFA, latent profile analyses (single-level and multilevel), multinomial logistic regression analysis	Australia and England	Lower and upper secondary education	Secondary data, Questionnaire survey to school leaders (principals) and teachers and student questionnaire	Job demands-resources theory, teacher wellbeing, latent profile analysis, multilevel, student achievement
13. Kang et al.	Original research	CFA, Structural equation modeling	China	Higher education	Questionnaire survey to students	School psychological capital, validation, academic engagement, achievement emotions, college students
14. Maulana et al.	Original research	Quantitative, CFA, multi-Group CFA (Measurement invariance)	England, Netherlands, the US	Secondary education	External observation by trained observers using an observation protocol	Classroom observation, measurement invariance, effective teaching behavior, 13 secondary education, cross-country comparison

10.3389/feduc.2023.1170854

TABLE 2 Specific components of the papers in relation to effective teaching.

Research Topics	Measurements	Antecedents	Correlates	Characteristics	Link with outcomes
1. Yang et al.	√		√	√	√
2. Li et al.				√	√
3. Ouwehand et al.		√	√		√
4. Gomariz et al.	√		√		
5. Melguizo-Garín et al.	√			√	√
6. Johansson et al.		√	√		
7. Pan et al.				√	
8. Liu		√	√		√
9. Gencoglu et al.		√	√	√	√
10. Falcón-Linares et al.			√	√	√
11. van de Pers and Helms-Lotenz		√		√	√
12. Collie et al.	√	√	√	√	√
13. Kang et al.	√		√		√
14. Maulana et al.	√		√		√

University students are often assigned to work on group projects as part of their courses. Developing competencies in group work skills is considered a more dynamic learning process closer to social reality, and thus, it is indispensable for their competitiveness in different workplaces and future career development. Using multiple regression, Melguizo-Garin et al. established a positive link between project-based learning (PBL) in university students by measuring competencies related to group work and satisfaction with PBL, suggesting the importance for university teachers to pay attention to the usefulness and planning of group work and forming group norms for a successful implementation of PBL.

Referred to as applying games or related elements, concepts, mechanisms or designs to classroom teaching, Game-based Learning (GBL) has become popular in practice and research when educational games are designed to promote self-regulated learning and enrich the classroom environment. The systematic literature review by Pan et al. is distinctive in its methodology to include an expert survey using the Delphi method, in addition to a comprehensive literature review, to develop a technology-enhanced GBL model which specifies the objectives, learning process, and evaluation of the approach in smart classroom environments.

While student perceptions of teaching quality are widely used as indicators of such quality in practice and research, variability exists in class and across cultures. Gencoglu et al. attempt to provide a conceptual framework to account for individual and cultural variations. By drawing on the relationships between cultural values, student perceptions and social desirability, the authors argue that their conceptual framework may help explain the variability in student perceptions and suggest that enhancing preservice teachers' awareness of value orientations in the classroom should be incorporated into teacher-training.

The paper of Maulana et al. examines the comparability of measuring effective teaching behavior using the International Comparative Analysis of Learning and Teaching (ICALT) observation instrument, and discusses how observed teaching behavior differs in three educational contexts (Netherlands, England, USA). Categorical multi-group confirmatory factor analysis was used. This study further identifies the need to understand the affordances and limitations of using an effective teaching measure to compare teaching quality across education systems.

2.2. Effective teaching, including factors beyond the classroom processes

Seven articles present studies on effective teaching which include factors beyond the classroom processes. These papers discuss various aspects of effective teaching exclusively, providing rich information about teachers' work that extend beyond instructional and pedagogical role in the classroom.

The paper by Gomariz et al. discusses the creation and validation of a teacher questionnaire called the "Questionnaire on Family Involvement in School, Teacher-Promoter" (QFIS-TP) for measuring teacher facilitation of family participation in Spain, including Early Childhood, Primary, and Secondary Education. They employed exploratory and confirmatory factor analyses as the primary analysis techniques and showed that the psychometric quality of the questionnaire was satisfactory. The study offers an additional tool for studying teachers' roles outside the classroom relating to engaging family involvement in schools.

Assuming some positive relationship between in-service professional teacher development and preservice teacher training is reasonable, but empirical evidence requires extensive data collection. Based on the analysis of national data of language and mathematics teachers in Sweden and multiple regression and multilevel growth modeling, Johansson et al.

showed that teachers with higher credits during teacher training were more likely to engage and perform better in professional development. Their study confirms Shulman's prediction that higher content knowledge contributes to more potent and more persistent development of pedagogical content knowledge.

The bachelor-level educational credentials of teachers are often argued to be indicative of effective teaching and contribute to pupils' learning outcomes. Conducting pupil fixed-effect modeling on the China Educational Panel Survey dataset, Liu provided additional evidence for the significance of teachers studying at degree level for students' learning gain. The effect of the educational credentials of teachers was estimated at about 1 month of additional learning per year. This study highlights the importance of educational credentials of teachers and suggests that learning gains can be improved by educating teachers more appropriately.

The Falcón-Linares et al. paper explores teachers' work in terms of the tutoring practice in Spanish secondary education. A semi-structured questionnaire was used as part of a mixed methods grounded theory approach. The study indicates a relationship between a teacher's profile and tutoring performance.

The authors conclude that, although teachers engage in coordination and supervision roles of tutoring, workloads and lack of clarity regarding tutoring norms create barriers to practicing effective tutoring.

The paper by van de Pers and Helms-Lorenz analyzes the relationship between secondary school background characteristics and beginning teachers' perceptions of induction activities in the Netherlands. Special attention is given to differences between professional development schools (PDS) and non-professional development schools (non-PDS). In a repeated measures design, questionnaires were used to measure perceptions of induction, support and linear regression analysis was performed. The authors found that multiple school locations, school size, the number of beginning teachers employed, and schools with a higher proportion of older male teachers are associated negatively with the frequency of induction activities. This study provides insight into enhancing equity in schools based on induction support perspectives.

Teachers' working conditions are essential for teachers' productive functioning. Investigations of job demands and resources can contribute to our understanding of how support for teachers can be improved. Using the TALIS 2018 data, the paper by Collie et al. investigates teachers' workplace experiences and their links with instructional support and achievement in Australia and England using a person-centered approach based on the job demand-resources model. Latent-profile analysis, multinomial logistic regression, and structural equation modeling were performed. The study reveals six teacher profiles and two school profiles. SES was found to link negatively with unsupportive school profiles. Instructional support and achievement were positively related to supportive school profiles.

Finally, the Kang et al. paper examines the validation of the school psychological capital (PsyCap) and its links with academic engagement and achievement emotions among university students in China. A student questionnaire was used, and structural equation modeling was employed. The paper demonstrates that the PsyCap questionnaire indicates a satisfactory psychometric quality. The authors found that perceived PsyCap resources are related positively to perceived academic engagement and enjoyment but negatively to perceived anxiety and boredom. Their study highlights the importance of school psychological capital resources for positive student outcomes.

3. Concluding thought

This Research Topic demonstrates how important it is to take a closer look at teachers' work with regard to effective teaching within and outside the classroom, considering issues related to measurements, antecedents, correlates, characteristics, and links with outcomes. All the papers included in this Research Topic contribute outstandingly to broadening our understanding of the complex nature of teaching and effectiveness. We thank the authors for their extraordinary contributions. We also thank all of the reviewers, as well as the Frontiers team for assuring the quality of scientific works. We hope that this collection of papers will stimulate international discussion and inspire further work to improve the knowledge base on effective teaching.

Author contributions

RM conceptualized and drafted the editorial paper and the Research Topic, wrote the paper and led the writing process, and was responsible for taking the role as first author. AK and JK wrote the paper, edited the draft, and were responsible for taking the role as co-authors. All authors contributed to the article and approved the submitted version.

Acknowledgments

We thank the authors for their extraordinary contributions. We also thank all of the reviewers, as well as the Frontiers team for assuring the quality of scientific works.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher,

the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ko, J., Sammons, P., and Bakkum, L. (2014). *Effective teaching* (Report No. 978-1-909437-75-3). Education Development Trust.

Maulana, R., Helms-Lorenz, M., and Klassen, R. (2023a). *Effective Teaching Around the World: Theoretical, Empirical, Methodological, and Practical Insights*. Dordrecht: Springer.

Maulana, R., Helms-Lorenz, M., Moorer, P., Smale-Jacobse, A., and Feng, X. (2023b). *Differentiated Instruction in Teaching From the International Perspective: Methodological and Empirical Insights*. Groningen: University of Groningen Press. doi: 10.21827/62c554175 9973

frontiers in Psychology

ORIGINAL RESEARCH
published: 09 July 2021
doi: 10.3389/fpsyg.2021.697703

Validation and Prediction of the School Psychological Capital Among Chinese College Students

Xia Kang[1], Yajun Wu[2] and Lisheng Li[2]*

[1]Teacher Education and Learning Leadership Unit, Faculty of Education, The University of Hong Kong, Hong Kong, China, [2]School of Foreign Languages and Literature, Yunnan Normal University, Kunming, China

This study validated the school psychological capital (PsyCap) scale in the Chinese context and examined the predictive effect of PsyCap resources on academic engagement and achievement emotions. Self-report data for PsyCap resources, student engagement, enjoyment, anxiety, and boredom toward English learning were collected from 1,000 sophomores. Item-level analyses and confirmatory factor analysis were used to verify the validity of the school PsyCap scale, and structural equation modeling was applied to reveal the predictive effect of school PsyCap resources on academic engagement and achievement emotions. Results showed that the school PsyCap scale retained superior psychometric properties. Besides, PsyCap resources were demonstrated to have a positive relationship to academic engagement and enjoyment, and a negative relationship to anxiety and boredom. The effectiveness of the school PsyCap scale was verified among Chinese college students, and besides the traditional predictors, school PsyCap is also critically important for students' academic engagement and achievement emotions. Limitations and implications are discussed.

Keywords: school psychological capital, validation, academic engagement, achievement emotions, college students

OPEN ACCESS

Edited by:
Ridwan Maulana,
University of Groningen, Netherlands

Reviewed by:
Sammy K. Ho,
Hong Kong Baptist University, China
Lidia Scifo,
Libera Università Maria SS. Assunta, Italy

****Correspondence:***
Xia Kang
xkang@hku.hk

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *20 April 2021*
Accepted: *18 June 2021*
Published: *09 July 2021*

Citation:
Kang X, Wu Y and Li L (2021) Validation and Prediction of the School Psychological Capital Among Chinese College Students. Front. Psychol. 12:697703. doi: 10.3389/fpsyg.2021.697703

INTRODUCTION

With the rise of positive psychology in educational research, the school PsyCap with positive psychology as the core has attracted more and more attention from scholars. The notion of PsyCap was firstly proposed in the industrial-organizational contexts to act as the crucial success factor for improving competitiveness and work efficiency (Luthans et al., 2004). Recently, the beneficial impact of PsyCap on academic outcomes (e.g., engagement, achievement emotions, subjective wellbeing, and academic performance) has also been valued in educational settings (Li et al., 2014; Siu et al., 2014; Datu et al., 2018; Carmona-Halty et al., 2019b; King et al., 2020). Psychological captial refers to the positive psychological state or psychological resources displayed in the process of individual growth and development, which is composed of four components of hope, optimism, resilience, and self-efficacy (Luthans et al., 2007b).

Although PsyCap has a significant positive effect on academic outcomes, notably, past studies mainly conducted in western contexts (Carmona-Halty et al., 2019a,b; Martínez et al., 2019), except for Li et al. (2014) study, few studies have focused on the PsyCap of Mainland Chinese students. Besides, prior studies mainly treated PsyCap construct as an observed variable without

mitigating the potential measurement errors (Li et al., 2014; Datu and Valdez, 2016). That is, attention should be paid to the synergies among the four subscales of PsyCap rather than merely treating them as independent variables. Under the circumstance that educators are increasingly aware of the importance of building positive psychological resources for students, Chinese students' PsyCap and its positive effects on achievement emotions and academic engagement need to be further explored.

Taking Chinese university students as participants, the present study aimed to fill these gaps by verifying the validity of the PsyCap scale and the potential synergies among the four subscales in Chinese university context, and exploring whether PsyCap has a predictive effect on the achievement emotions and academic engagement.

LITERATURE REVIEW

School Psychological Capital

Psychological captial concentrates on an individual's positive psychological capacities and advantages. Since the notion of PsyCap was first introduced into the field of education, the characteristics of students' PsyCap and its relationship with academic outcomes have received educators and practitioners' extensive attention. However, as PsyCap is a formative measurement model constructed base on the synchronicity of the academic community, its essential characteristics, components, and synergies among components are all controversial.

There are three views on the essential characteristics of PsyCap, namely, trait-like, state-like, and bundling. Hosen et al. (2003) and Peterson and Seligman (2004) regarded PsyCap as a stable internal psychological structure produced by individual self-investment. By considering the developability of PsyCap, Luthans (2002) and Luthans et al. (2007b) argued that PsyCap is a state-like attitude rather than psychological traits. The bundling view of PsyCap holds that PsyCap has both state-like and trait-like aspects (Luthans et al., 2005). Given that all the four components of PsyCap (i.e., self-efficacy, hope, optimism, and resilience) are state-like positive psychological forces, and PsyCap can be cultivated *via* training, PsyCap was usually regarded as malleable and state-like (Luthans et al., 2005, 2007a).

Self-efficacy (or confidence), hope, optimism, and resilience are the four psychological capacities that contribute to the formation of positive PsyCap. Self-efficacy refers to an individual's conviction that he or she has the ability to execute challenging tasks. As a key psychological resource, self-efficacy can effectively alleviate the negative impact of stress on individuals (e.g., Luthans, 2007b) and has a positive predictive effect on academic performance (Galla et al., 2014; Jiang et al., 2014), optimistic achievement emotions (Putwain et al., 2018), and student motivation (Bong and Clark, 1999; Jiang et al., 2014). Snyder (2000, p. 9) proposed that hope is goal-related thinking, which consists of three components, namely, "goals, pathways, and agency". The positive predictive effects of hope have been verified in existing studies. For example, Snyder et al. (2002) and Day et al. (2010) took college students as participants and found that hope had a positive association with academic success. Moreover, the contributions of hope to positive emotions (Aspinwall and Leaf, 2002; Khodarahimi, 2013) and psychological adjustment (Du and King, 2013) were also confirmed. From the perspective of attribution theory, Seligman (2002) regarded optimism as the explanatory style for handling events: permanence and pervasiveness. Specifically, the psychological capacity of optimism refers to individuals' expectation of positive outcomes for their endeavors. The predictive effects of optimism were also vastly explored in the academic context. For instance, optimism has been positively linked to academic performance (Hoy et al., 2006; Feldman and Kubota, 2015; Icekson et al., 2020) and academic engagement (Medlin and Faulk, 2011). Resilience refers to the capacity of sustaining and bouncing back when encountering problems and adversities (Luthans et al., 2007a). That is, students high in resilience are more likely to view adversities as challenges rather than threats (Symes et al., 2015). The positive effects of resilience on academic outcomes have also been widely explored. For example, studies have shown that academic resilience was the positive predictor of classroom participation, enjoyment of school, self-esteem (Martin and Marsh, 2006), academic buoyancy (Martin and Marsh, 2009), academic performance (Kotzé and Kleynhans, 2013; Ayala and Manzano, 2018), and psychological health (García-Izquierdo et al., 2018). To sum up, each component of PsyCap has important and positive significance for students' academic outcomes.

Measurement of School Psychological Capital

Diachronically, both the constructs of PsyCap and the interaction mode of these constructs were in the process of dynamic development. Until recent years, there has been a consensus on the constructs of PsyCap within the academic community; however, the higher-order PsyCap instrument in the academic context remains to be explored.

In the early days, the self-esteem scale was used to measure a Person's PsyCap, that is, the only construct of PsyCap was self-esteem (Goldsmith et al., 1997). Afterward, Letcher and Niehoff (2004) equalized PsyCap to five constructs, those are neuroticism, extraversion, openness, agreeableness, and conscientiousness. Later, Avey et al. (2006) and Larson and Luthans (2006) used four constructs (i.e., self-efficacy, hope, resilience, and optimism) to represent the PsyCap measure. Specifically, the four PsyCap components are composed of Scheier and Carver's (1985) optimism scale, Wagnild and Young's (1993) resilience scale, Snyder et al.'s (1996) hope scale, and Parker's (1998) self-efficacy scale. So far, the use of the four psychological capacities of self-efficacy, hope, resilience, and optimism to characterize PsyCap has become a consensus of the academic circles (Luthans et al., 2007b), and its application scope has also begun to expand from the industrial-organizational contexts to the school contexts.

Except for King and Caleon's (2021) study, very few studies have explored the psychometric properties of the PsyCap scale

in school contexts. King and Caleon (2021) firstly adapted the items in Luthans et al.'s (2007a) PsyCap scale into school-related ones, and then tested the validity of the school PsyCap scale with Singapore secondary students as participants. The overall Cronbach's alpha reliability of the school PsyCap scale was $\alpha = 0.91$, and the model fit of the scale was adequate. However, the school PsyCap scale developed by King and Caleon (2021) was only verified in the Singapore context and this questionnaire was presented in English, less is known about the validity of this scale in non-English-speaking countries, especially China.

According to the constitutional forms of the items, four competing models of school PsyCap scale coexist, namely, the null model, the unidimensional model, the four-factor model, and the hierarchical model. For example, Li et al. (2014) obtained the value of PsyCap by summing-up and averaging the values of the items, PsyCap was regarded as an omnibus construct, but the measurement error was ignored. Datu and Valdez (2016) and Carmona-Halty et al. (2019b) posited the school PsyCap scale as a unidimensional structure, that is, the four components of the PsyCap scale were treated as observational variables. Recently, Datu et al. (2018) and King et al. (2020) viewed school PsyCap structure as a hierarchical model, that is, both PsyCap and its four components were posited as latent constructs. In the Singapore context, King and Caleon (2021) conducted confirmatory factor analyses to examine the fit of the unidimensional model, the four-factor model, and the hierarchical model of the school PsyCap scale and found that the hierarchical model was the most acceptable one. Given the coexistence of multiple models of school PsyCap scale, more research is needed to examine which model has the best fit.

The Present Study

The present study had two objectives. The first one was to examine the construct validity of the school PsyCap scale among the Chinese population. Specifically, the psychometric properties and factor structure of the school PsyCap scale were tested with Chinese college students as participants. The second objective was to examine whether school PsyCap was the predictor of the outcome variables of academic engagement and achievement emotions among Chinese college students.

Both in industrial-organizational contexts and school contexts, the previous studies argued that the PsyCap scale has the best fitting degree when it was regarded as a hierarchical model (Luthans et al., 2007a; Datu et al., 2018; King and Caleon, 2021). The hierarchical model posits PsyCap as a second-order latent variable with self-efficacy, hope, resilience, and optimism as its first-order latent factors. Accordingly, we formed the first hypothesis of this study.

> *H1*: The hierarchical model of the school PsyCap scale would have the best model fit in the Chinese university context. More clearly, this study posited that school PsyCap would be a second-order latent factor with self-efficacy, resilience, hope, and optimism as the first-order latent factors.

Prior studies both in organizational and educational contexts have indicated that PsyCap is positively correlated with adaptive outcomes and negatively correlated with maladaptive outcomes. For example, the predictive effects of PsyCap on adaptive outcomes, such as subjective wellbeing (Li et al., 2014; Datu and Valdez, 2016), academic engagement (King et al., 2020), intrinsic motivation (Siu et al., 2014), academic satisfaction (Ortega-Maldonado and Salanova, 2018), and academic performance (Datu et al., 2018; Carmona-Halty et al., 2019a,b; Martínez et al., 2019) were confirmed. On the other hand, the negative effects of PsyCap on maladaptive outcomes, such as academic procrastination (Hicks and Wu, 2015), substance abuse problems (Krasikova et al., 2015), and depressive symptoms (King and Caleon, 2021) were also confirmed. The control-value theory, as the framework for studying achievement emotions, holds that control and value appraisals are the two most important antecedents of individuals' achievement emotions (Pekrun et al., 2002; Pekrun, 2006, 2009). Besides, considering that PsyCap is more stable than emotions (Luthans et al., 2007b) and individuals with higher PsyCap are more likely to experience positive emotions (Avey et al., 2011), it can be posited that PsyCap would be the predictor of achievement emotions in school contexts. Correspondingly, we formed the second hypothesis of this study.

> *H2*: School PsyCap would be positively correlated with positive achievement emotions, such as enjoyment, and negatively related to negative achievement emotions, such as anxiety and boredom.

Besides, the job demand-resources theory holds that individual resources can not only positively affect the level of job engagement (Akanni et al., 2019), but also change the job demands into challenges (Siu et al., 2014). Since PsyCap is a kind of individual resource (e.g., economic capital, social capital, and human capital), we form the third hypothesis of this study.

> *H3*: School PsyCap would have a positive predictive effect on academic engagement among Chinese college students.

MATERIALS AND METHODS

Participants and Procedure

A total of 1,000 sophomores who learn English as a foreign language at a university in southwest China participated in this study. The mean age of the participants was 19.43 years (SD = 0.83 years). The sample comprised 215 males (21.5%) and 785 females (78.5%). The gender distribution in the sample is consistent with the gender ratio of normal universities in Mainland China.

The scales used in this study were firstly translated from English into Chinese, and then back-translated by bilingual researchers to ensure the equivalence of the Chinese version of

these scales. The questionnaire consists of the school PsyCap scale, engagement and disaffection scale, and achievement emotion scale, which took about 25 min to complete. Participants completed the questionnaire in an English class under the guidance of their English teachers.

Measures

School Psychological Capital

The Chinese version of the school PsyCap scale was adapted from King and Caleon's (2021) school psychological capital scale. The school psychological capital scale is in English, and it was adapted from the psychological capital scale that was originally used to measure employees' PsyCap capacity in industrial-organizational contexts (Luthans et al., 2007a). The important contribution of King and Caleon's (2021) study was that the items in the psychological capital scale were replaced and adapted for the first time to be relevant to the school context. In their study (King and Caleon, 2021), five experts in psychology and education were asked to evaluate and classify the 24 items on the PsyCap scale and those items that received 80% of the experts' approval would be retained. Finally, 16 items were retained to measure students' school PsyCap capacity. The lowest acceptable factor loading is 0.40 (Matsunaga, 2010), and the factor loading of one item of the optimism scale in the present study (i.e., "Overall, I expect more good things than bad things to happen to me in school") was below this criteria. Therefore, we adapted and translated King and Caleon's (2021) 16-item school psychological capital scale into a 15-item Chinese school PsyCap scale.

This school PsyCap scale asks students to mark their agreement or disagreement with the descriptions, and the items were rated by a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree). Sample items include as follows: "I feel confident that I can learn what is taught in school" (four items; self-efficacy); "if I have problems in school, I could think of many ways to solve them" (four items; hope); "I think I'm good at dealing with schoolwork pressures" (four items; resilience); and "I always look on the positive side of things in school" (three items; optimism).

Academic Engagement

Skinner and colleague's engagement and disaffection scale (Skinner et al., 2009) were adapted to measure students' behavioral and emotional engagement. The behavioral engagement subscale included four items, such as "I try hard to do well in English class," and the emotional engagement subscale included four items, such as "English class is fun." The cognitive engagement subscale was adapted from Reeve and Tseng's (2011) Engagement Scale, an sample item is "when I study, I try to connect what I am learning with my own experiences." The cognitive engagement subscale also consisted of four items. All three engagement subscales were rated on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

In line with the previous studies (e.g., Lam et al., 2014), this study also treated engagement as a hierarchical construct, that is, academic engagement was posited as a second-order latent factor underpinned by the first-order latent factors of behavioral, emotional, and cognitive engagement. The Cronbach's alpha reliability of the total engagement scale was 0.94. The Cronbach's alpha reliability of each subscale of engagement was also examined, and the results showed that all the subscales had good reliability indices: behavioral engagement (α = 0.79), emotional engagement (α = 0.82), and cognitive engagement (α = 0.88). Moreover, the results of confirmatory factor analyses indicated the hierarchical model had excellent fit: (χ^2 = 233.745; df = 51; p < 0.001; χ^2/df = 4.58; RMSEA = 0.060 [90% CI: 0.052–0.068]; SRMR = 0.029; CFI = 0.971; and TLI = 0.963), which supported the hierarchical model of academic engagement.

Achievement Emotions

This study considered three concrete achievement emotions, namely, enjoyment, boredom, and anxiety because they are most frequently and intensely experienced by students during learning activities. According to Pekrun's classification, enjoyment is perceived as positive emotions, while anxiety and boredom are classified as negative emotions (Pekrun, 2006; Pekrun and Linnenbrink-Garcia, 2012). We adapted items from Pekrun et al.'s (2011) achievement emotions questionnaire to measure students' achievement emotions. Specifically, four items were used each to measure enjoyment, anxiety, and boredom. Answers were anchored on a 7-point Liker scale ranging from 1 (strongly disagree) to 7 (strongly agree). All subscales had good internal reliability: α = 0.837 for enjoyment, α = 0.835 for anxiety, and α = 0.763 for boredom.

Statistical Analyses

We executed detailed item analyses to examine the psychometric properties of the school PsyCap scale. After this, Confirmatory Factor Analysis (CFA) models were conducted to examine the structure of the school PsyCap. We tested three models, namely, the unidimensional model, four-factor model, and hierarchical model, to determine which model fits the data well. A series of indices were used to evaluate model fit: root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square error of approximation (SRMR). The values of CFI and TLI more than 0.90 mean that the model fit is adequate, and the values of RMSEA and SRMR less than 0.08 are regarded as acceptable (Hu and Bentler, 1999; Chen, 2007). Afterward, multigroup second-order CFA model was used to evaluate the measurement invariance of school PsyCap scale across gender groups.

To assess between-network validity, the zero-order correlations between school PsyCap and a wide range of academic outcomes, such as academic engagement, enjoyment, anxiety, and boredom, were tested. Subsequently, structural equation model (SEM) was conducted with PsyCap as predictor and engagement, enjoyment, anxiety, and boredom as outcome variables. Gender and age were considered as covariates.

RESULTS

Treatment of Missing Data

This study used questionnaires of 1,000 Chinese sophomores as the analytical sample, and 0.4% of the questionnaires had missing responses. The expectation-maximization technique was

applied to supplement the loss of information caused by missing responses.

Psychometric Properties of the School PsyCap Scale

Item-Level Analyses

A set of item analyses were firstly conducted to assess the item quality of the school PsyCap scale. Precisely, the means and variance, the distributional properties, the corrected item-total correlation, and Cronbach's alpha if item deleted of all the 15 items were analyzed and evaluated. The results of item-level analyses for the school PsyCap scale are shown in **Table 1**.

Each item of the school PsyCap scale satisfies a normal distribution for both the skewness and kurtosis value conformed to the criteria suggested by Finney and DiStefano (2006), and the value of corrected item-total correlation of these items met the criteria proposed by Clark and Watson (1995). The skewness of all the items ranges from −0.04 to −0.51, which is within the criteria range of −2 to +2. The kurtosis values of all the items range from −0.76 to 0.04, which is also within the criteria range of −7 to +7. In addition, the values of corrected item-total correlations range from $r = 0.50$ to $r = 0.74$, which satisfy the criteria of $r > 0.40$.

The reliability (Cronbach's alpha) of the total school PsyCap scale is $\alpha = 0.91$. In the present study, the Cronbach's alpha for each of the subscales of school PsyCap measure was 0.81 (self-efficacy), 0.73 (hope), 0.78 (resilience), and 0.68 (optimism). The results indicated that the total school PsyCap scale had good reliability and the reliability of the four subscales was also acceptable.

Subsequently, the four subscales' Cronbach's alpha reliability were checked again *via* item deletion. That is, we inspected whether the subscale's Cronbach's alpha reliability would increase or decrease by deleting one of its items. For instance, we looked at whether deleting the item "if I have problems in school, I could think of many ways to solve them" will increase or decrease the Cronbach's alpha for hope. As shown in **Table 1**, we found that the 15 items of the school PsyCap scale would only result in lower Cronbach's alpha reliability when any item deletion was conducted.

Confirmatory Factor Analysis

To further examine the psychometric properties of the school PsyCap scale, a series of confirmatory factor analyses were conducted. Specifically, three competing models were used to fit the data. Model 1 was a unidimensional model which posited that all the 15 items loaded into an omnibus school PsyCap construct. Model 2 was a four-factor model that had four inter-correlated latent constructs: self-efficacy, hope, resilience, and optimism. Model 3 was a hierarchical model that posited PsyCap as a second-order latent factor underpinned by the four first-order latent factors of self-efficacy, hope, resilience, and optimism.

As shown in **Table 2**, the model fit of the unidimensional model (model 1) was grudgingly acceptable. Both the

TABLE 1 | Item-level analyses for the school PsyCap scale.

Item	Mean	*SD*	Skewness	Kurtosis	Corrected item-total correlation	Cronbach's alpha if item deleted
EFF1	4.75	1.28	−0.51	−0.04	0.65	0.76
EFF2	4.58	1.31	−0.20	−0.24	0.61	0.75
EFF3	4.30	1.41	−0.16	−0.50	0.61	0.79
EFF4	4.61	1.19	−0.30	0.04	0.74	0.77
HO1	4.88	1.16	−0.31	−0.25	0.62	0.68
HO2	4.55	1.28	−0.37	−0.01	0.57	0.65
HO3	3.80	1.36	−0.04	−0.35	0.57	0.67
HO4	4.45	1.34	−0.18	−0.39	0.63	0.67
RES1	4.71	1.19	−0.25	−0.10	0.70	0.74
RES2	5.05	1.29	−0.51	−0.24	0.53	0.74
RES3	4.54	1.24	−0.24	−0.27	0.72	0.69
RES4	4.26	1.50	−0.05	−0.76	0.56	0.76
OPT1	4.93	1.27	−0.32	−0.35	0.50	0.64
OPT2	4.75	1.23	−0.32	−0.32	0.61	0.54
OPT3	4.30	1.43	−0.17	−0.44	0.62	0.57

EFF, self-efficacy; HO, hope; RES, resilience; and OPT, optimism.

TABLE 2 | CFA results of the three competing models.

Model	χ^2	*df*	χ^2/df	RMSEA	90% CI	CFI	TLI	SRMR
Model 1	587.108***	90	6.52	0.074	0.069, 0.080	0.920	0.907	0.040
Model 2	430.270***	84	5.12	0.064	0.058, 0.070	0.944	0.930	0.035
Model 3	476.006***	86	5.53	0.067	0.062, 0.073	0.937	0.923	0.037

***$p < 0.001$.

four-factor model (model 2) and the hierarchical model (model 3) had an adequate fit. The target coefficient, which refers to the ratio of the chi-square of the first-order model to the chi-square of the higher-order model, is used to prove the existence of a higher-order construct and the upper limit of the target coefficient is 1 (Marsh and Hocevar, 1985). The target coefficient of the school PsyCap scale in this study is 0.90, indicating that this scale is a higher-order construct. Coupled with theoretical reasons (King and Caleon, 2021), the hierarchical model was adopted.

Measurement Invariance of School PsyCap Across Gender Groups

To examine the invariance of school PsyCap scale across gender groups, multigroup second-order CFA model was conducted. Followed by the principles of testing invariance in a second-order CFA model recommend by Wang and colleagues (Wang and Wang, 2019), this study firstly tested second-order configural invariance of the school PsyCap scale, and then tested the invariance of second-order factor loadings. Measurement invariance will be established if the following two requirements are satisfied: (1) the overall model fit is adequate (Little, 1997) and (2) the value of ΔCFI between two nested models should be smaller than or equal to 0.01 (Cheung and Rensvold, 2002).

Second-Order Configural CFA Model

In the present study, school PsyCap was viewed as a second-order construct with self-efficacy, hope, resilience, and optimism as the first-order latent factors. In the second-order configural CFA model, the measurement parameters and structural parameters can be freely estimated, and covariances between the residual terms of the first-order factors were all set to zero. Besides, for the purpose of model identification, intercepts of the first-order factors and the means of the second-order factors were all also set to zero.

The results showed that the second-order configural model fits the data well: $\chi^2 = 611.640$; $df = 171$; $p < 0.001$; $\chi^2/df = 3.58$; RMSEA = 0.072 [90% CI: 0.066–0.078]; SRMR = 0.041; CFI = 0.929; and TLI = 0.913. This indicated that the results of the configural model can be used as the baseline values against which the specified restricted models can be compared.

Testing Invariance of Second-Order Factor Loadings

The invariance of first-order factor loadings and item intercepts are the prerequisites for verifying the measurement invariance of the second-order factor loadings. Thus, both the invariance of the first-order factor loadings and item intercepts across gender groups would be firstly identified.

As shown in **Table 3**, the overall model fit was good and all ΔCFIs were smaller than 0.01, indicating that configural, metric, and scalar invariances were established in first-order factors of the school PsyCap.

After checking the invariance of first-order factor loadings and the item intercepts, we tested the invariance of the second-order factor loadings by way of testing whether the relations between the four first-order factors (i.e., self-efficacy, hope, resilience, and optimism) and school PsyCap are invariant across gender groups. Specifically, we imposed equality restrictions on both first-order factors and second-order factors, and then compared model fit between the current model and the second-order configural CFA model. The model fit of current model was good: $\chi^2 = 689.938$; $df = 200$; $p < 0.001$; $\chi^2/df = 3.45$; RMSEA = 0.070 [90% CI: 0.065–0.076]; SRMR = 0.060; CFI = 0.921; and TLI = 0.917. Comparing with the second-order configural CFA model, we got the following results: ΔCFI = 0.929–0.921 = 0.008 < 0.01. That is, the second-order factor loadings of school PsyCap were invariant across gender groups.

School PsyCap as a Predictor of Academic Outcomes

The psychometric properties of the school PsyCap scale were sufficient for use among Chinese college students and confirmatory factor analysis indicated that the school PsyCap is best treated as a second-order latent variable. After this, we aimed to examine whether school PsyCap is positively linked to some key forms of academic outcomes.

Bivariate Correlations

The zero-order correlations between school PsyCap and several academic outcomes were examined. More specifically, we focused on the correlations between school PsyCap and academic engagement and three discrete achievement emotions (enjoyment, anxiety, and boredom). As expected, school PsyCap was positively related to academic engagement and positive emotions (i.e., enjoyment) and negatively correlated with negative emotions (i.e., anxiety and boredom). The results are presented in **Table 4**.

Structural Equation Modeling

We constructed a SEM in which school PsyCap is an exogenous variable and academic engagement, enjoyment, anxiety, and boredom are outcome variables (see **Figure 1**). The model fit was good: $\chi^2 = 2299.760$; $df = 755$; $p < 0.001$; $\chi^2/df = 3.05$; RMSEA = 0.050 [90% CI: 0.047–0.052]; SRMR = 0.047; CFI = 0.912; and TLI = 0.905, which indicates that the proposed model fits the sample well. Results are showed in **Figure 1**. In line with our hypotheses, school PsyCap has positive predictive effect on academic engagement ($\beta = 0.53$, $p < 0.001$) and positive emotion (i.e., enjoyment; $\beta = 0.46$, $p < 0.001$). However, the predictive effect of school PsyCap on negative emotions is negative, particularly, anxiety ($\beta = -0.30$, $p < 0.001$) and boredom ($\beta = -0.25$, $p < 0.001$).

DISCUSSION

In the present study, we adapted and translated the existing school PsyCap scale (King and Caleon, 2021) into Chinese. Then, the psychometric properties of the 15-item school PsyCap scale were evaluated with Chinese college students as participants. Next, the association between school PsyCap and adaptive and

TABLE 3 | Testing first-order factor loadings and item intercepts across genders.

Model	χ^2	*df*	CFI	ΔCFI	TLI	RMSEA	90% CI	SRMR
M1: Configural invariance	556.923	167	0.937	–	0.921	0.069	0.063, 0.075	0.039
M2: Metric invariance	582.338	178	0.935	0.002	0.923	0.068	0.062, 0.074	0.048
M3: Scalar invariance	633.885	193	0.929	0.006	0.923	0.068	0.062, 0.074	0.056

TABLE 4 | Results of descriptive, bivariate correlations, and Cronbach's alpha reliability.

S. No.		1	2	3	4	5	6	7
1.	School PsyCap	–	0.482**	0.401**	−0.236**	−0.200**	−0.079*	0.038
2.	Engagement		–	0.826**	−0.444**	−0.561**	0.130**	0.012
3.	Enjoyment			–	−0.511**	−0.596**	0.156**	−0.007
4.	Anxiety				–	0.747**	−0.110**	0.006
5.	Boredom					–	−0.202**	−0.003
6.	Gender						–	−0.139**
7.	Age							–
	Mean	4.57	4.51	4.44	3.61	3.34	1.79	19.41
	SD	0.87	0.91	1.07	1.26	1.22	0.41	1.01
	Cronbach's alpha	0.91	0.94	0.84	0.84	0.76	–	–

*$p < 0.05$; **$p < 0.01$.

maladaptive learning-related outcomes were examined. We found that the 15-item school PsyCap scale had good psychometric properties and the measurement invariance of this scale was also confirmed in the Chinese university context and that school PsyCap has a significant predictive effect on academic engagement and achievement emotions. This study contributes to the literature on PsyCap which has mainly been confined to the industrial-organization or western cultural contexts but has not yet appeared in the Chinese context.

By comparing three competing models of the school PsyCap scale, school PsyCap was suggested to be regarded as a higher-order latent variable underpinned by the four first-order latent variables of self-efficacy, hope, resilience, and optimism. Hypothesis one was supported. This result is consistent with prior studies on PsyCap in industrial-organization contexts (Larson and Luthans, 2006), Philippine context (Datu et al., 2018), and Singapore context (King and Caleon, 2021). In addition to the theoretical reasons proposed by King and Caleon (2021), the present study also evaluated the target coefficient of the school PsyCap scale to support the hierarchical conceptualization of school PsyCap. The four components of school PsyCap have extensively been explored in the educational context, and the significance of the present study lies in the validation that school PsyCap is a second-order latent variable with its four components as first-order latent variables.

The present study also mined both the positive and negative associations between school PsyCap and achievement emotions. Precisely, we found that school PsyCap has a positive predictive effect on positive achievement emotions, such as enjoyment. School PsyCap also negatively predicted negative achievement emotions, such as boredom and anxiety. The second hypothesis of this study was also confirmed. These findings were consistent with the previous studies that school PsyCap was positively correlated with adaptive outcomes, such as subjective wellbeing (Li et al., 2014), interdependent happiness, flourishing (Datu and Valdez, 2016), positive affect, and life satisfaction (King and Caleon, 2021). On the other hand, our findings were also consistent with prior studies that PsyCap negatively predicted maladaptive outcomes, such as emotional exhaustion, depersonalization (Ding et al., 2015), procrastination (Hicks and Wu, 2015), and depressive symptoms (King and Caleon, 2021). To our knowledge, this study is the first to explore and verify the association between school PsyCap and achievement emotions among Chinese college students. Achievement emotions are crucial to learning (Zull, 2006; Tyng et al., 2017), and it is of great value to promote college students' achievement emotions by way of intervening their school PsyCap.

We also found that school PsyCap is the positive predictor of academic engagement, which fully supports the third hypothesis. Although the referents of academic engagement are not identical, the consistent finding of this study and the previous studies is that school PsyCap has a positive predictive effect on academic engagement. Taking Pilipino high school students as participants, Datu and Valdez (2016) utilized behavioral and emotional engagement to represent academic engagement and confirmed that PsyCap had positive predictive effect on academic engagement. Our finding is also consistent with King and Caleon's (2021) study that school PsyCap positively predicated academic engagement. In King and Caleon's (2021) study, the scores of behavioral, emotional, and cognitive engagement were aggregated to form an overall score of academic emotions. Academic engagement is one of the optimal academic-related outcomes, and exploring the link between school PsyCap and academic engagement would have positive implications for the enhancement of academic outcomes.

Overall, the present study contributes to the literature on school PsyCap by (1) validating the school PsyCap scale among

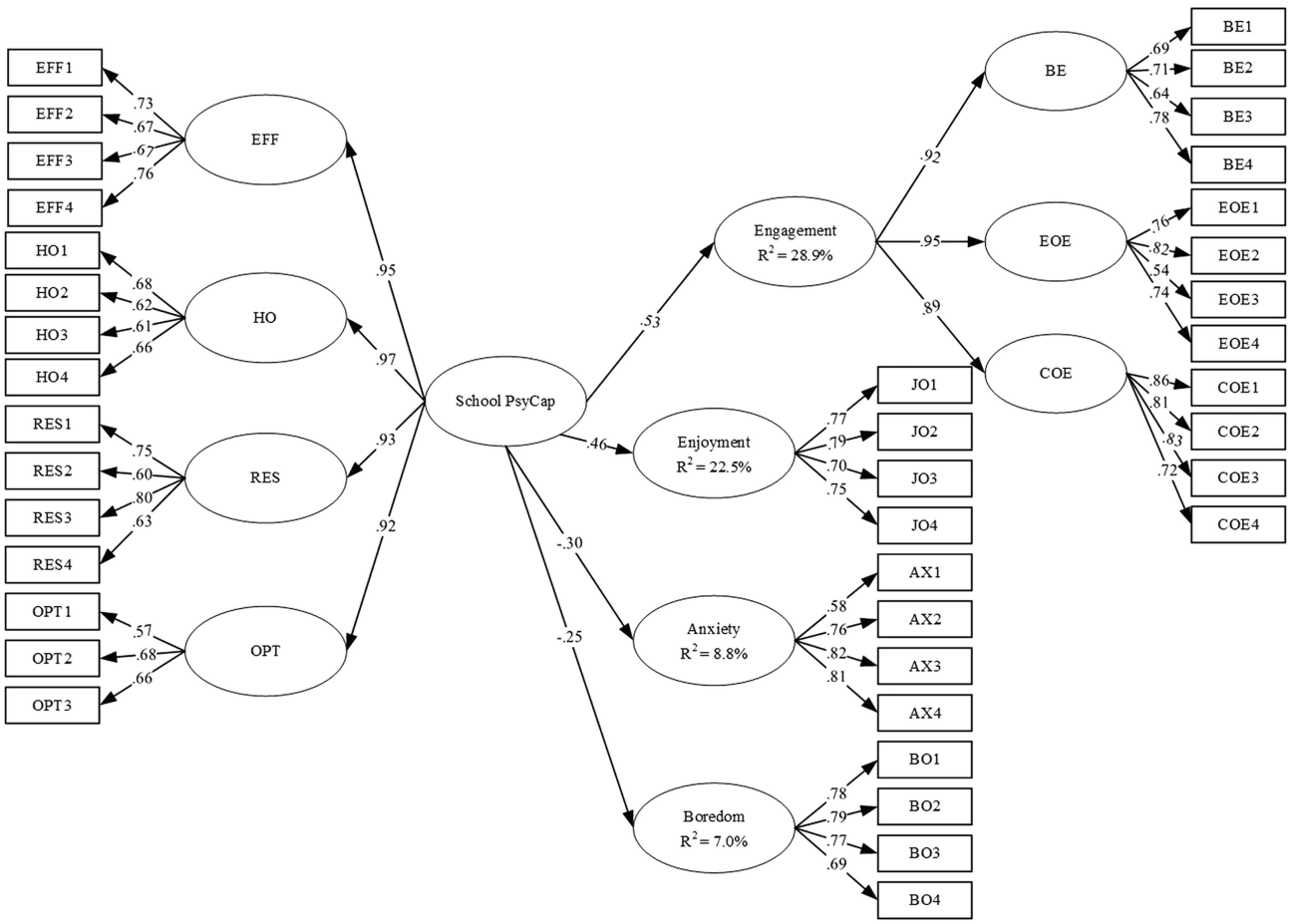


FIGURE 1 | Final structural equation model on school PsyCap and academic outcomes. EFF, self-efficacy; HO, hope; RES, resilience; OPT, optimism; BE, behavioral engagement; EOE, emotional engagement; COE, cognitive engagement. All parameters in the figure are standardized coefficient and all paths are significant at $p < 0.001$. For brevity reason, the residuals are not presented in the figure. Gender and age are the covariates.

Chinese college students and demonstrating that school PsyCap is best viewed as a higher-order variable, (2) confirming both positive and negative predictive effects of school PsyCap on achievement emotions, and (3) verifying the positive association between school PsyCap and academic engagement. Since previous PsyCap-related studies were mostly carried out in western contexts (e.g., Carmona-Halty et al., 2019b,c; Martínez et al., 2019), this study took Chinese college students as participants to explore the effectiveness of school PsyCap scale as well as the predictive effects of school PsyCap on achievement emotions and academic engagement. The findings of our study indicated that school PsyCap may also be an important resource in some non-western cultural contexts, such as the Chinese university context.

LIMITATION AND DIRECTIONS FOR FUTURE RESEARCH

Although this study expanded the application scope of the school PsyCap scale and explored the correlations between school PsyCap and achievement emotions and academic engagement, four limitations need to be addressed. Firstly, data for all the measures were self-reported and the risk of common method bias cannot be completely avoided. In addition to the self-reported data, future studies are suggested to collect data from teachers and peers to reduce the common method bias. Secondly, samples from more colleges/universities are needed to make the research more representative. Chinese colleges and universities recruit students from all provinces in China, that is, every single university includes college students that come from all provinces of China. It is a typical situation that there are more female students than male students in normal universities in China and the correlations between appraisal antecedents and achievement emotions are equivalent across genders (Pekrun, 2018). Nevertheless, the female students took a high proportion in the present sample, which calls for future studies to select more male students to balance the gender distribution. Thirdly, the present study only took Chinese college students as the survey sample. Although both the psychometric properties and the target coefficient of the school PsyCap scale were evaluated with Chinese college students as

participants, future studies should include Chinese elementary and secondary school students to further broaden the application scope of the school PsyCap scale. Lastly, this study was correlational and thus causal conclusions cannot be drawn. Given that the main purpose was to demonstrate the validity of the school PsyCap scale in the Chinese context, the correlational nature of this study was a necessity. However, the relations among school PsyCap, achievement emotions, and academic engagement might be dynamic reciprocal, and thus cross-lagged panel design is suggested for future studies.

EDUCATIONAL IMPLICATIONS

The validation of the 15-item school PsyCap scale indicates that college administrators and teachers can utilize this scale to assess college students' school PsyCap capacity in Chinese settings. Also, the significant correlations between school PsyCap and achievement emotions and academic engagement suggest that nurturing students' school PsyCap capacity might be one appropriate way to enhance their academic and wellbeing outcomes. Furthermore, the psychological capital is more measurable and malleable than the traditional human and social capital (Luthans et al., 2004), which implicates that the enhancement of students' PsyCap capacity could be one more effective way for the development of their academic outcomes.

Although the research on promoting students' PsyCap capacity has not yet appeared extensively (Carmona-Halty et al., 2019b), existing studies have shown that it was possible to develop a student's PsyCap capacity by promoting the four components of PsyCap (Luthans et al., 2008, 2010). Firstly, relating today's learning to tomorrow's life is one possible way for teachers to enhance their students' PsyCap capacity because high-hope individuals would have clear goals, replenish their willpower, and feel excited about their future (Marques et al., 2017). Secondly, teachers and educators are suggested to provide clear expectations for students (Eley and Stallman, 2014) as well as increase students' control and choice over their studies (Brewer et al., 2019) to enhance students' ability to bounce back when facing challenges or adversities. Thirdly, instructors can change the self-evaluation of those stressful students by, for example, listing solutions to their perceived stress to maintain an optimistic outlook about their studies (Rand et al., 2020). Fourthly, self-efficacy was believed to be the dominating component of human agency and teachers are the credible resources to nourish students' self-efficacy beliefs by way of persuading and providing positive feedback so that students can experience mastery in learning (Usher and Pajares, 2008; Honicke and Broadbent, 2016).

The affective issues of learning activities are practically relevant to effective teaching (Kyriacou, 2009). Empirical studies show that students' academic success is positively correlated with their positive achievement emotions and negatively correlated with the negative achievement emotions (e.g., Putwain et al., 2013; Dewaele et al., 2018), which indicates that increasing students' positive achievement emotions (e.g., enjoyment) and reducing their negative achievement emotions (e.g., anxiety and boredom) would be the effective methods for teachers to motivate students to sustain further efforts in learning activities. The present study confirmed the correlation between school PsyCap and achievement emotions among Chinese college students; therefore, teachers and educators are suggested to influence on students' achievement emotions by way of improving their psychological capital. In addition, the correlation between school PsyCap and academic engagement was also verified, which indicates that the enhancement of students' psychological capital would contribute to effective teaching because students' involvement and engrossment in learning activities would also be increased in this process (Fredricks et al., 2004).

CONCLUSION

Given the significance of school PsyCap to education, the present study is the first to demonstrate the validity of the school PsyCap scale with Chinese college students as participants. Results also indicate that school PsyCap is a higher-order latent variable underpinned by first-order latent variables of self-efficacy, hope, optimism, and resilience. Besides, the predictive effects of school PsyCap on achievement emotions and academic engagement were also confirmed. Compared with other forms of capital (e.g., economic and social capital), it is more cost-effective to improve academic and wellbeing outcomes by developing the school PsyCap. Therefore, educators and teachers are suggested to nurture students' school PsyCap capacity by creating environments and developing interventions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Human Research Ethics Committee of the University of Hong Kong. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

XK designed, analyzed the survey, and wrote the manuscript. YW conducted the survey and edited the manuscript. LL proofread the manuscript. All authors contributed to the article and approved the submitted version.

REFERENCES

Akanni, A. A., Kareem, D. B., and Oduaran, C. A. (2019). Emotional intelligence and work engagement among bank workers: moderated mediation model of ethical leadership and job burnout. *Int. J. Work Organ. Emot.* 10, 322–338. doi: 10.1504/IJWOE.2019.106885

Aspinwall, L. G., and Leaf, S. L. (2002). In search of the unique aspects of hope: pinning our hopes on positive emotions, future-oriented thinking, hard times, and other people. *Psychol. Inq.* 13, 276–288. doi: 10.1207/S15327965PLI1304_02

Avey, J. B., Patera, J. L., and West, B. J. (2006). The implications of positive psychological capital on employee absenteeism. *J. Leadersh. Org. Stud.* 13, 42–60. doi: 10.1177/10717919070130020401

Avey, J. B., Wernsing, T. S., and Mhatre, K. H. (2011). A longitudinal analysis of positive psychological constructs and emotions on stress, anxiety, and wellbeing. *J. Leadersh. Org. Stud.* 18, 216–228. doi: 10.1177/1548051810397368

Ayala, J. C., and Manzano, G. (2018). Academic performance of first-year university students: the influence of resilience and engagement. *High. Educ. Res. Dev.* 37, 1321–1335. doi: 10.1080/07294360.2018.1502258

Bong, M., and Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educ. Psychol.* 34, 139–153. doi: 10.1207/s15326985ep3403_1

Brewer, M. L., van Kessel, G., Sanderson, B., Naumann, F., Lane, M., Reubenson, A., et al. (2019). Resilience in higher education students: a scoping review. *High. Educ. Res. Dev.* 38, 1105–1120. doi: 10.1080/07294360.2019.1626810

Carmona-Halty, M., Salanova, M., Llorens, S., and Schaufeli, W. B. (2019a). How psychological capital mediates between study-related positive emotions and academic performance. *J. Happiness Stud.* 20, 605–617. doi: 10.1007/s10902-018-9963-5

Carmona-Halty, M., Salanova, M., Llorens, S., and Schaufeli, W. B. (2019b). Linking positive emotions and academic performance: the mediated role of academic psychological capital and academic engagement. *Curr. Psychol.*, 40, 2938–2947. doi: 10.1007/s12144-019-00227-8

Carmona-Halty, M., Schaufeli, W. B., and Salanova, M. (2019c). Good relationships, good performance: the mediating role of psychological capital - A three-wave study among students. *Front. Psychol.* 10:306. doi: 10.3389/fpsyg.2019.00306

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* 14, 464–504. doi: 10.1080/10705510701301834

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Model.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Clark, L. A., and Watson, D. (1995). Construct validity: basic issues in objective scale development. *Psychol. Assess.* 7, 309–319.

Datu, J. A. D., King, R. B., and Valdez, J. P. M. (2018). Psychological capital bolsters motivation, engagement, and achievement: cross-sectional and longitudinal studies. *J. Posit. Psychol.* 13, 260–270. doi: 10.1080/17439760.2016.1257056

Datu, J. A. D., and Valdez, J. P. M. (2016). Psychological capital predicts academic engagement and wellbeing in Filipino high school students. *Asia-Pacific Edu. Res.* 25, 399–405. doi: 10.1007/s40299-015-0254-1

Day, L., Hanson, K., Maltby, J., Proctor, C., and Wood, A. (2010). Hope uniquely predicts objective academic achievement above intelligence, personality, and previous academic achievement. *J. Res. Pers.* 44, 550–553. doi: 10.1016/j.jrp.2010.05.009

Dewaele, J. M., Witney, J., Saito, K., and Dewaele, L. (2018). Foreign language enjoyment and anxiety: the effect of teacher and learner variables. *Lang. Teach. Res.* 22, 676–697. doi: 10.1177/1362168817692161

Ding, Y., Yang, Y., Yang, X., Zhang, T., Qiu, X., He, X., et al. (2015). The mediating role of coping style in the relationship between psychological capital and burnout among Chinese nurses. *PLoS One* 10:e0122128. doi: 10.1371/journal.pone.0122128

Du, H., and King, R. B. (2013). Placing hope in self and others: exploring the relationships among self-construals, locus of hope, and adjustment. *Personal. Individ. Differ.* 54, 332–337. doi: 10.1016/j.paid.2012.09.015

Eley, D. S., and Stallman, H. (2014). Where does medical education stand in nurturing the 3Rs in medical students: responsibility, resilience and resolve? *Med. Teach.* 36, 835–837. doi: 10.3109/0142159X.2014.917159

Feldman, D. B., and Kubota, M. (2015). Hope, self-efficacy, optimism, and academic achievement: distinguishing constructs and levels of specificity in predicting college grade-point average. *Learn. Individ. Differ.* 37, 210–216. doi: 10.1016/j.lindif.2014.11.022

Finney, S. J., and DiStefano, C. (2006). "Non-normal and categorical data in structural equation models" in *A Second Course in Structural Equation Modeling*. eds. G. R. Hancock and R. O. Mueller (Greenwich: Information Age), 269–314.

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059

Galla, B. M., Wood, J. J., Tsukayama, E., Har, K., Chiu, A. W., and Langer, D. A. (2014). A longitudinal multilevel model analysis of the within-person and between-person effect of effortful engagement and academic self-efficacy on academic performance. *J. Sch. Psychol.* 52, 295–308. doi: 10.1016/j.jsp.2014.04.001

García-Izquierdo, M., Ríos-Risquez, M. I., Carrillo-García, C., and Sabuco-Tebar, E. D. L. Á. (2018). The moderating role of resilience in the relationship between academic burnout and the perception of psychological health in nursing students. *Educ. Psychol.* 38, 1068–1079. doi: 10.1080/01443410.2017.1383073

Goldsmith, A. H., Veum, J. R., and Darity, W. (1997). The impact of psychological and human capital on wages. *Econ. Inq.* 35, 815–829. doi: 10.1111/j.1465-7295.1997.tb01966.x

Hicks, R., and Wu, F. M. Y. (2015). Psychological capital as mediator between adaptive perfectionism and academic procrastination. *GSTF J. Psychol.* 2, 34–40. doi: 10.5176/2345-7872

Honicke, T., and Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: a systematic review. *Educ. Res. Rev.* 17, 63–84. doi: 10.1016/j.edurev.2015.11.002

Hosen, R., Solovey-Hosen, D., and Stern, L. (2003). Education and capital development: capital as durable personal, social, economic and political influences on the happiness of individuals. *Education* 123, 496–514.

Hoy, W. K., Tarter, C. J., and Woolfolk Hoy, A. (2006). Academic optimism of schools: a force for student achievement. *Am. Educ. Res. J.* 43, 425–446. doi: 10.3102/00028312043003425

Hu, L.-T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6, 1–55. doi: 10.1080/10705519909540118

Icekson, T., Kaplan, O., and Slobodin, O. (2020). Does optimism predict academic performance? Exploring the moderating roles of conscientiousness and gender. *Stud. High. Educ.* 45, 635–647. doi: 10.1080/03075079.2018.1564257

Jiang, Y., Song, J., Lee, M., and Bong, M. (2014). Self-efficacy and achievement goals as motivational links between perceived contexts and achievement. *Educ. Psychol.* 34, 92–117. doi: 10.1080/01443410.2013.863831

Khodarahimi, S. (2013). Hope and flourishing in an Iranian adults sample: their contributions to the positive and negative emotions. *Appl. Res. Qual. Life* 8, 361–372. doi: 10.1007/s11482-012-9192-8

King, R. B., and Caleon, I. S. (2021). School psychological capital: instrument development, validation, and prediction. *Child Indic. Res.* 14, 341–367. doi: 10.1007/s12187-020-09757-1

King, R. B., Pitliya, R. J., and Datu, J. A. (2020). Psychological capital drives optimal engagement via positive emotions in work and school contexts. *Asian J. Soc. Psychol.* 23, 457–468. doi: 10.1111/ajsp.12421

Kotzé, M., and Kleynhans, R. (2013). Psychological wellbeing and resilience as predictors of first-year students' academic performance. *J. Psychol. Afr.* 23, 51–59. doi: 10.1080/14330237.2013.10820593

Krasikova, D. V., Lester, P. B., and Harms, P. D. (2015). Effects of psychological capital on mental health and substance abuse. *J. Leadersh. Org. Stud.* 22, 280–291. doi: 10.1177/1548051815585853

Kyriacou, C. (2009). *Effective Teaching in Schools: Theory and Practice*. Cheltenham: Nelson Thornes.

Lam, S., Jimerson, S., Shin, H., Cefai, C., Veiga, F. H., Hatzichristou, C., et al. (2014). Understanding and measuring student engagement in school: the results of an international study from 12 countries. *Sch. Psychol. Q.* 29, 213–232. doi: 10.1037/spq0000057

Larson, M., and Luthans, F. (2006). Potential added value of psychological capital in predicting work attitudes. *J. Leadersh. Org. Stud.* 13, 45–62. doi: 10.1177/10717919070130010701

Letcher, L., and Niehoff, B. (2004). *Psychological Capital and Wages: A Behavioral Economic Approach*, Minneapolis: Kansas State University.

Li, B., Ma, H., Guo, Y., Xu, F., Yu, F., and Zhou, Z. (2014). Positive psychological capital: a new approach to social support and subjective wellbeing. *Soc. Behav. Pers.* 42, 135–144. doi: 10.2224/sbp.2014.42.1.135

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: practical and theoretical issues. *Multivar. Behav. Res.* 32, 53–76. doi: 10.1207/s15327906mbr3201_3

Luthans, F. (2002). The need for and meaning of positive organizational behavior. *J. Organ. Behav.* 23, 695–706. doi: 10.1002/job.165

Luthans, F., Avey, J. B., Avolio, B. J., and Peterson, S. J. (2010). The development and resulting performance impact of positive psychological capital. *Hum. Resour. Dev. Q.* 21, 41–66. doi: 10.1002/hrdq.20034

Luthans, F., Avey, J. B., and Patera, J. L. (2008). Experimental analysis of a web-based training intervention to develop positive psychological capital. *Acad. Manag. Learn. Educ.* 7, 209–221. doi: 10.5465/amle.2008.32712618

Luthans, F., Avolio, B. J., Avey, J. B., and Norman, S. M. (2007a). Positive psychological capital: measurement and relationship with performance and satisfaction. *Pers. Psychol.* 60, 541–572. doi: 10.1111/j.1744-6570.2007.00083.x

Luthans, F., Avolio, B. J., Walumbwa, F. O., and Li, W. (2005). The psychological capital of Chinese workers: exploring the relationship with performance. *Manag. Organ. Rev.* 1, 249–271. doi: 10.1111/j.1740-8784.2005.00011.x

Luthans, F., Luthans, K. W., and Luthans, B. C. (2004). Positive psychological capital: beyond human and social capital. *Bus. Horiz.* 47, 45–50. doi: 10.1016/j.bushor.2003.11.007

Luthans, F., Youssef, C. M., and Avolio, B. J. (2007b). *Psychological Capital: Developing the Human Competitive Edge*. New York: Oxford Univerisity Press.

Marques, S. C., Gallagher, M. W., and Lopez, S. J. (2017). Hope-and academic-related outcomes: a meta-analysis. *Sch. Ment. Heal.* 9, 250–262. doi: 10.1007/s12310-017-9212-9

Marsh, H. W., and Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: first- and higher order factor models and their invariance across groups. *Psychol. Bull.* 97, 562–582. doi: 10.1037/0033-2909.97.3.562

Martin, A. J., and Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: a construct validity approach. *Psychol. Sch.* 43, 267–281. doi: 10.1002/pits.20149

Martin, A. J., and Marsh, H. W. (2009). Academic resilience and academic buoyancy: multidimensional and hierarchical conceptual framing of causes, correlates and cognate constructs. *Oxf. Rev. Educ.* 35, 353–370. doi: 10.1080/03054980902934639

Martínez, I. M., Youssef-Morgan, C. M., Chambel, M. J., and Marques-Pinto, A. (2019). Antecedents of academic performance of university students: academic engagement and psychological capital resources. *Educ. Psychol.* 39, 1047–1067. doi: 10.1080/01443410.2019.1623382

Matsunaga, M. (2010). How to factor-analyze your data right: do's, dont's, and how-to's. *Int. J. Psychol. Res.*. 3, 97–110. doi: 10.21500/20112084.854

Medlin, B., and Faulk, L. (2011). The relationship between optimism and engagement: the impact on student performance. *Res. High. Educ. J.* 13.

Ortega-Maldonado, A., and Salanova, M. (2018). Psychological capital and performance among undergraduate students: the role of meaning-focused coping and satisfaction. *Teach. High. Educ.* 23, 390–402. doi: 10.1080/13562517.2017.1391199

Parker, S. K. (1998). Enhancing role breadth self-efficacy: the roles of job enrichment and other organizational interventions. *J. Appl. Psychol.* 83, 835–852. doi: 10.1037/0021-9010.83.6.835

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educ. Psychol. Rev.* 18, 315–341. doi: 10.1007/s10648-006-9029-9

Pekrun, R. (2009). "Emotions at school," in *Handbook of Motivation at School*. eds. K. R. Wentzel and A. Wigfield (New York and London: Routledge), 575–604.

Pekrun, R. (2018). "Control-value theory: A social-cognitive approach to achievement emotions," in *Big Theories Revisited 2: A Volume of Research on Sociocultural Influences on Motivation and Learning*. eds. G. A. D. Liem and D. M. McInerney (Charlotte: Information Age Publishing), 162–190.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., and Perry, R. P. (2011). Measuring emotions in students' learning and performance: the Achievement Emotions Questionnaire (AEQ). *Contemp. Educ. Psychol.* 36, 36–48. doi: 10.1016/j.cedpsych.2010.10.002

Pekrun, R., Goetz, T., Titz, W., and Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ. Psychol.* 37, 91–105. doi: 10.1207/S15326985EP3702_4

Pekrun, R., and Linnenbrink-Garcia, L. (2012). "Academic emotions and student engagement" in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York: Springer), 259–282.

Peterson, C., and Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Oxford: Oxford Univerisity Press.

Putwain, D. W., Pekrun, R., Nicholson, L. J., Symes, W., Becker, S., and Marsh, H. W. (2018). Control-value appraisals, enjoyment, and boredom in mathematics: a longitudinal latent interaction analysis. *Am. Educ. Res. J.* 55, 1339–1368. doi: 10.3102/0002831218786689

Putwain, D., Sander, P., and Larkin, D. (2013). Academic self-efficacy in study-related skills and behaviours: relations with learning-related emotions and academic success. *Br. J. Educ. Psychol.* 83, 633–650. doi: 10.1111/j.2044-8279.2012.02084.x

Rand, K. L., Shanahan, M. L., Fischer, I. C., and Fortney, S. K. (2020). Hope and optimism as predictors of academic performance and subjective wellbeing in college students. *Learn. Individ. Differ.* 81:101906. doi: 10.1016/j.lindif.2020.101906

Reeve, J., and Tseng, C. M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemp. Educ. Psychol.* 36, 257–267. doi: 10.1016/j.cedpsych.2011.05.002

Scheier, M., and Carver, C. S. (1985). Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health Psychol.* 4, 219–247. doi: 10.1037/0278-6133.4.3.219

Seligman, M. E. P. (2002). *Authentic Happiness*. New York: The Free Press.

Siu, O. L., Bakker, A. B., and Jiang, X. (2014). Psychological capital among university students: relationships with study engagement and intrinsic motivation. *J. Happiness Stud.* 15, 979–994. doi: 10.1007/s10902-013-9459-2

Skinner, E. A., Kindermann, T. A., and Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educ. Psychol. Meas.* 69, 493–525. doi: 10.1177/0013164408323233

Snyder, C. R. (2000). *Handbook of Hope: Theory, Measures, and Applications*. San Diego: Academic Press.

Snyder, C. R., Shorey, H. S., Cheavens, J., Pulvers, K. M., Adams, V. H., and Wiklund, C. (2002). Hope and academic success in college. *J. Educ. Psychol.* 94, 820–826. doi: 10.1037/0022-0663.94.4.820

Snyder, C. R., Sympson, S. C., Ybasco, F. C., Borders, T. F., Babyak, M. A., and Higgins, R. L. (1996). Development and validation of the state hope scale. *J. Pers. Soc. Psychol.* 70, 321–335. doi: 10.1037/0022-3514.70.2.321

Symes, W., Putwain, D. W., and Remedios, R. (2015). The enabling and protective role of academic buoyancy in the appraisal of fear appeals used prior to high stakes examinations. *Sch. Psychol. Int.* 36, 605–619. doi: 10.1177/0143034315610622

Tyng, C. M., Amin, H. U., Saad, M. N. M., and Malik, A. S. (2017). The influences of emotion on learning and memory. *Front. Psychol.* 8:1454. doi: 10.3389/fpsyg.2017.01454

Usher, E. L., and Pajares, F. (2008). Sources of self-efficacy in school: critical review of the literature and future directions. *Rev. Educ. Res.* 78, 751–796. doi: 10.3102/0034654308321456

Wagnild, G. M., and Young, H. M. (1993). Development and psychometric evaluation of the resilience scale. *J. Nurs. Manag.* 1, 165–178.

Wang, J., and Wang, X. (2019). *Structural Equation Modeling: Applications Using Mplus*. West Sussex: John Wiley and Sons.

Zull, J. E. (2006). "Key aspects of how the brain learns," in *The Neuroscience of Adult Learning*. eds. S. Johnson and K. Taylor (San Francisco: Jossey-Bass), 3–9.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 06 August 2021
doi: 10.3389/fpsyg.2021.711173

A Multilevel Person-Centered Examination of Teachers' Workplace Experiences: Replication and Extension With Links to Instructional Support and Achievement

Rebecca J. Collie[1]*, *Andrew J. Martin*[1], *Alexandre J. S. Morin*[2], *Lars-Erik Malmberg*[3] and *Pamela Sammons*[3]

[1] School of Education, University of New South Wales, Sydney, NSW, Australia, [2] Substantive Methodological Synergy Research Laboratory, Department of Psychology, Concordia University, Montreal, QC, Canada, [3] Department of Education, University of Oxford, Oxford, United Kingdom

In a replication and extension of an earlier study, we relied on person-centered analyses to identify teacher (Level 1) and school (Level 2) profiles based on teachers' experiences of job demands (barriers to professional development, disruptive student behavior), job resources (teacher collaboration, input in decision-making), and personal resources (self-efficacy). We examined data from 5,439 teachers working in 364 schools in Australia and 2,216 teachers working in 149 schools in England. Latent profile analysis revealed six teacher profiles: Low-Demand-Flourisher (11%), Mixed-Demand-Flourisher (17%), Job-Resourced-Average (11%), Balanced-Average (14%), Mixed-Resourced-Struggler (11%), and Low-Resourced-Struggler (36%). Two school profiles were identified: an Unsupportive school profile (43%) and a Supportive school profile (57%). Several significant relations between these profiles and teacher/school characteristics and work-related outcomes were also identified at both levels. Although our results generally replicated prior findings, some differences were also observed, possibly as a results of recent changes in policies regarding in teacher support and accountability. Next, we extended prior work using a subsample of the Australian teachers for whom we had matching student data. This second set of results revealed that schools with a greater proportion of low-SES students were more likely to present an Unsupportive school profile. Moreover, the Supportive school profile was associated with higher levels of student-reported instructional support and school-average achievement in reading, mathematics, and science.

Keywords: job demands-resources theory, teacher well-being, latent profile analysis, multilevel, student achievement

OPEN ACCESS

Edited by:
Ridwan Maulana,
University of Groningen, Netherlands

Reviewed by:
Antti-Tuomas Pulkka,
National Defence University, Finland
Leon T. de Beer,
North-West University, South Africa

***Correspondence:**
Rebecca J. Collie
rebecca.collie@unsw.edu.au
orcid.org/0000-0001-9944-2703

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *18 May 2021*
Accepted: *14 July 2021*
Published: *06 August 2021*

Citation:
Collie RJ, Martin AJ, Morin AJS, Malmberg L-E and Sammons P (2021) A Multilevel Person-Centered Examination of Teachers' Workplace Experiences: Replication and Extension With Links to Instructional Support and Achievement. Front. Psychol. 12:711173. doi: 10.3389/fpsyg.2021.711173

INTRODUCTION

Teachers' exposure to a variety of demands and resources at work is known to be associated with important workplace outcomes (e.g., Dicke et al., 2018; Skaalvik and Skaalvik, 2018). Recently, researchers have begun to examine demand-resource profiles to better understand the role played by distinct combinations of work-related demands and resources, and the extent to which these

combinations are linked to different outcomes among teachers (e.g., Simbula et al., 2012; Collie and Martin, 2017). Emerging work is also considering these profiles at the school-level. More precisely, researchers have started to examine whether different types of schools can be identified based on the prevalence of different teacher demand-resource profiles (Collie et al., 2020a). The combined identification of teacher and school profiles can be helpful for informing policy and practice to improve teachers' experiences at work, and potentially their students' outcomes as well.

Collie et al. (2020a) provided the first ever multilevel person-centered investigation of teacher and school demand-resource profiles, along with an examination of how these profiles were related to teacher- and school-level outcomes. They did so using data collected in the Teaching and Learning International Survey (TALIS) 2013 from Australia and England, which are two countries that have similarities in their educational systems and student populations (for a discussion, see Collie et al., 2020a; see also Fackler et al., 2020). The first goal of the present study is to assess whether Collie et al.'s (2020a) seminal research findings will be replicated among new samples of Australian and English teachers using data collected 5 years later as part of TALIS 2018, and relying on the same measures of job demands (barriers to professional development, disruptive student behavior), job resources (teacher collaboration, input in decision-making), and personal resources (self-efficacy). Given the many changes and increases in complexity that have impacted teachers' work between 2013 and 2018 (discussed in detail below; Guerriero and Révai, 2017), testing for replication using large-scale representative data appears to be critically important for at least two reasons. First, it provides a means to better understand teachers' contemporary experiences. Second, it provides a means to assess whether and how these recent changes may have affected teachers' work orientation. Thus, following from Collie et al. (2020a), the present study seeks to identify teacher and school demand-resource profiles separately in each country, before relying on systematic tests of profile similarity to assess replication across countries. We then consider how these profiles are related to teacher and school characteristics and to outcomes located at both levels of analysis (teacher and school-average job satisfaction and occupational commitment). **Figure 1** displays the models under examination.

The second goal of the current study is to extend Collie et al.'s (2020a) research (and other prior work) by examining whether and how the identified school profiles are associated with differences in student-reported outcomes. Examining the extent to which school profiles are linked with students' outcomes has the potential to guide intervention efforts aiming to promote healthy and effective schools—both for teachers and students. In the current study, we examined (school-average) student perceptions of instructional support provided by teachers (autonomy-support, instrumental help, and warmth) and achievement in reading, mathematics, and science. In this second phase, data from a subsample of Australian teachers was used, along with matched student data aggregated at the school-level from the Programme for International Student Assessment (PISA) 2018 study.

Conceptual Framework and Relevance for Person-Centered Approaches

We rely on the Job Demands-Resources (JD-R) theoretical model (Bakker and Demerouti, 2017) as the conceptual framework for the present study. According to the JD-R framework, all jobs are seen as involving a range of demands and resources that are psychological, social, physical or organizational in nature. Whereas, job demands (e.g., high workload) entail a cost for the employee, and hinder employee functioning at work, job resources (e.g., social support) represent a gain for the employee and help to foster employee positive functioning at work (Bakker and Demerouti, 2017). JD-R theory also establishes that employees bring with them their own personal resources (e.g., self-efficacy), which act as additional types of resources in fostering employee positive functioning at work. Thus, the JD-R theoretical framework assumes that job demands, job resources, and personal resources are all directly associated with important work-related outcomes for the exposed employees, such as motivation, well-being, and performance. In addition to those direct processes, JD-R theory proposes two interactive processes. In the first, which is called the buffering process, job resources are posited to reduce the detrimental impact of job demands on negative outcomes like burnout and stress. In the second, which is called the boosting process and was the focus in the present study, high job demands are posited to enhance the beneficial impact of resources on positive outcomes such as motivation and well-being (Bakker and Demerouti, 2017). The boosting process occurs because job resources become more useful and relevant when demands are high.

A growing body of work has examined teachers' experiences of demands and resources at work. These studies generally show that job demands tend to be associated with poorer work-related outcomes for the exposed teachers (e.g., greater emotional exhaustion, lower well-being, and reduced work commitment; Skaalvik and Skaalvik, 2018), whereas job and personal resources tend to be associated with more positive outcomes (e.g., greater job satisfaction, lower stress; Collie et al., 2018). In addition, research has provided support for the boosting process, whereby job resources appear to share a stronger association with positive outcomes when job demands are high (Bakker et al., 2007). Notably, however, this prior research has largely been conducted using variable-centered approaches (e.g., multiple regression, path analyses) that describe how isolated types of demands and resources are associated with outcomes on average. Because variable-centered approaches rest on the assumption of population-homogeneity, they provide information about associations at a sample-wide level and about particular variables that could be targeted in broad intervention efforts.

More recently, researchers have begun relying on person-centered (e.g., latent profile analysis; Mäkikangas et al., 2018) investigations of the JD-R theoretical framework. Person-centered approaches allow for the identification of different subpopulations (or profiles) of teachers characterized by qualitatively distinct configurations of demands and resources. In this context, person-centered approaches appear to be highly relevant to the development of intervention procedures

tailored to distinct profiles of teachers. With respect to JD-R theory, person-centered approaches are particularly relevant to investigating the boosting (and buffering) process (Collie et al., 2020a). Indeed, whereas variable-centered research is able to examine the boosting process *via* tests of interactions among pairs of variables, it is difficult to clearly interpret the results from analyses involving more than three interacting factors. In contrast, person-centered approaches are able to directly consider the combined role played by any number of demands and resources simultaneously. This has practical implications for schools, in which principals typically have to deal with "types" (or profiles) of teachers presenting different characteristics, rather than with the isolated and combined effects of variables. As such, person-centered profiles tend to be readily recognizable in practice. Importantly, person-centered analysis can be used to identify both individual- and school-level profiles, such as teacher and school profiles, making it possible to devise initiatives targeting specific types of teachers, as well as specific types of schools.

Overview of Prior Research on Demand-Resource Profiles

A small, but growing body of work has identified demand-resource profiles among employees and organizations, with a handful of these studies having more specifically focused on teachers and schools. For example, Simbula et al. (2012) investigated two job resources (professional development and collegial support) and two job demands (role ambiguity and over investment in work) among teachers. Their findings revealed three profiles: high demand-high resource, high demand-low resource, and low demand-high resource. Among an undifferentiated sample of employees, Van den Broeck et al. (2012) identified three similar profiles, along with a fourth (low resource-low demand) in a study focusing on several job demands (e.g., workload, emotional and cognitive demands) and job resources (e.g., social support, autonomy). In terms of organization-level profiles, Mäkikangas et al. (2018) examined department-level demand-resource profiles among university employees. They showed two employee profiles (low demand-high-resource vs. high demand-low resource), along with two department-level profiles: a high stress department (dominated by employees corresponding to the high demand-low resource profile) and a mixed-stress department (characterized a relatively even mix of the two employee profiles). More recently, Collie et al. (2020a) conducted a multi-level investigation of demand-resource profiles among teachers (Level 1) and schools (Level 2). Because this study forms the basis for the current study, we now consider their work in greater detail.

Key Findings From an Earlier Study

Collie et al. (2020a) examined five demands and resources that are known to play a major role in the capability of teachers to effectively undertake their work. *Barriers to professional development* is a first type of job demand referring to teachers' exposure to issues preventing them from accessing the training required for their ongoing learning and development (e.g., financial constraints, limited opportunities; OECD, 2009; Broadley, 2010). *Disruptive student behavior* is a second type of job demand referring to teachers' exposure to student off-task behavior that hinders effective instruction (e.g., students refusing to listen, antisocial behaviors; Skaalvik and Skaalvik, 2018). *Teacher collaboration* is a first type of job resource referring to the extent to which teachers are able to efficiently work with their colleagues to plan, develop, teach, and assess student learning (OECD, 2014). *Teacher input in decision-making* is a second type of job resource referring to teachers' perceptions that their school provides them with opportunities to have an input in, and share responsibility for, school decisions (OECD, 2014). Finally, *self-efficacy for teaching* is a personal resource referring to teachers' confidence in their ability to efficiently undertake a range of tasks designed to ensure and maximize student learning (Tschannen-Moran and Woolfolk Hoy, 2001). The present study focuses on the same demands and resources.

Taken together, these five demands and resources reflect relatively common experiences in teaching that impact teachers' ability to be effective and to thrive at work (e.g., Klassen et al., 2012). More precisely, these five factors encompass teachers' interactions with students, colleagues, and school leaders, as well as teachers' own professional development and confidence. Teachers who have positive interactions with others at school, and who feel confident in their abilities, have been found to be more likely to experience positive work-related outcomes (e.g., Klassen et al., 2012). The opposite is true for teachers who have challenging interactions with students or who do not experience agency in relation to their professional growth (e.g., Skaalvik and Skaalvik, 2018).

At the teacher-level (Level 1), Collie et al. (2020a) identified five demand-resource profiles that were equivalent across samples of teachers from Australia and England. The first profile, called the Low-Demand-Flourisher, was characterized by low demands and high resources (both job and personal). The second profile, called the Mixed-Demand-Flourisher, was characterized by low-average demands and high resources. The third profile, the Job-Resourced-Average profile, was characterized by average job demands, above average job resources, and average personal resources. The fourth profile, the Balanced-Average profile reported average levels across all demands and resources. Finally, the Struggler profile reported high demands and low resources. At the school-level (Level 2), Collie et al. (2020a) identified two school profiles: a Supportive school profile comprising relatively equal levels of the Mixed-Demand-Flourisher, Job-Resourced-Average, Struggler, and Low-Demand-Flourisher profiles, and an Unsupportive school profile comprising much higher levels of the Struggler profile. These school profiles were also found to be equivalent in Australia and England.

Summary

Taken together, prior research provides growing understanding of the nature of the demand-resource profiles most commonly identified among teachers and schools. At the same time, given the limited number of previous studies, more research is needed to ascertain the extent to which the same number of profiles, characterized by a similar shape, will be identified in new samples, and whether these profiles will display similar outcome

implications and associations with predictors. This verification is important in order to formally assess the extent to which previous results can be expected to generalize to new samples of participants, which forms an important pre-requisite to the development of intervention procedures guided by these person-centered results (e.g., Morin et al., 2016). Furthermore, it is also important to consider how these profiles will related to novel predictors and outcomes for extending our understanding of their nomological network.

As noted earlier, the aim of the present study is to replicate Collie et al.'s (2020a) research, but also to extend this earlier study. Replication was undertaken by conducting person-centered analyses on the same set of demands and resources among a similar sample of Australian and English teachers (Level 1) and schools (Level 2) using data collected 5 years after those examined in the Collie et al. (2020a) study. In the upcoming pages, we first outline how changes in policies and practices occurring between 2013 and 2018 might have impacted teachers' work experiences over this time. These changes provide a compelling case for the importance of conducting a replication study. We then more specifically address the various predictors and outcomes considered in the present study, as well as how this study was designed to extend upon Collie et al.'s (2020a) results.

Recent Changes in Teachers' Work

Although many aspects of teachers' work roles remained similar between 2013 and 2018, there was also (inevitable) ongoing evolution of the teaching profession that resulted in some changes to teachers' experiences of demands and resources at work. Worldwide, there has recently been concern about the growing complexity of the teaching profession and the rising demands being placed upon teachers (e.g., Guerriero and Révai, 2017; OECD, 2019a). In both Australia and England, significant concerns about high workload have been raised by governments (e.g., UK Department for Education, 2018; Parliament of Australia, 2019). Alongside these concerns, there were also policy shifts in Australia and England between 2013 and 2018 that may have impacted teachers' perceptions of work. Thus, both countries implemented new policy priorities toward improving teacher quality, including increases to school funding and the promotion of teachers' professional development and collaboration (OECD, 2019a). This may have led to greater access to professional development and higher levels of collaboration with colleagues. At the same time, (student-reported) disruptive student behavior has recently been found to be higher in Australia and England than in many other OECD countries (OECD, 2019a), and was found to have increased in Australia over the past decade (Thomson et al., 2020). Increases in compliance, accountability, and external evaluation are also apparent in both countries between 2013 and 2018 (OECD, 2019a). These changes might have meant that teachers experienced increases in disruptive behavior and decreases in their ability to provide input in decision-making. Compliance and accountability are also known to be associated with reduced self-efficacy among teachers (von der Embse et al., 2016).

Taken together, although teachers' work roles remained broadly similar from 2013 to 2018, policies and practices changed during this period in a way that might have impacted teachers' experiences, and demand-resource profiles, at work. For example, increases in accountability and compliance (OECD, 2019a) might be accompanied by shifts in teachers' experiences of input and self-efficacy at work, as well as in the way these two types of resources tend to be associated for teachers. Alternatively, if we assume that the identified profiles reflect more enduring differences among teachers, who work in an occupation that is known to experience changes in policies and practices in an ongoing manner, then these changes might not have impacted the nature of the observed demand-resource profiles. Arguably, these two possibilities are both likely. Thus, although we do expect the identified profiles to generalize (which is a critical condition to our ability to use them to guide interventions that will be efficient over time), we leave as an open research question whether this generalizability will be complete, or partial. Furthermore, it is also possible that the changes occurring between 2013 and 2018 might have modified the associations between the observed profiles and the predictors and outcomes. We now turn our attention to these predictors and outcomes.

Teacher and School Characteristics That Predict Profile Membership

Teacher background characteristics have been shown to predict membership in demand-resource profiles. For example, male teachers and more experienced teachers tend to correspond to less adaptive teacher profiles (e.g., Collie et al., 2020a). Given these results, teacher sex and teaching experience were examined as predictors of the teacher profiles identified in the present study. Little is known, however, in relation to variables involved in the prediction of school profiles. For instance, Mäkikangas et al. (2018) found no associations between departmental size and the department-level profiles identified among university employees. Notably, Collie et al. (2020a) did not examine school-level predictors in their study. In the present study, school-level predictors were thus considered as a way to expand upon Collie et al.'s (2020a) prior work. In light of prior variable-centered research results highlighting the key role played by various school characteristics, this inclusion is deemed to be important. For instance, different school locations are known to come with unique supports and challenges that can impact teachers' workplace experiences (Klassen and Chiu, 2010). For example, schools in less populous locations and with lower socio-economic status typically have less access to teaching resources and professional development opportunities (e.g., Broadley, 2010). Larger schools have been shown to engage in more teacher collaboration (Collie et al., 2020b), but to have lower levels of teacher input (Collie, 2021). Likewise, in English-speaking countries, schools with more students from non-English speaking backgrounds have been shown to have lower levels of teacher collaboration (Collie et al., 2020b). Given these variable-centered associations, it appeared important to consider the extent to which various school characteristics are also associated with school demand-resource profiles.

In sum, the present study first provided an opportunity to verify whether Collie et al.'s (2020a) results regarding the role

played by two teacher characteristics (sex, teaching experience) in predicting teachers' membership into the various profiles would generalize in this new sample collected 5 years later. In addition, it also allowed us to expand on these previous results by also considering the role played by four school characteristics (school location, school size, proportion of students from low socio-economic status [SES] backgrounds, and proportion of students from non-English speaking backgrounds) in the prediction of the school-level profiles. We anticipated that the findings obtained at the teacher-level would be similar to those reported by Collie et al. (2020a). At the school-level, based on the previously reported variable-centered results, we hypothesized that schools located in more populous locations, as well as schools with fewer students from low-SES and non-English speaking backgrounds would be more likely to display a Supportive profile. Given mixed findings related to school size, we left the nature of the associations between this school characteristic and the profiles as an open research question.

Teacher and School Profiles and Associations With Teacher and Student Outcomes

Membership in demand-resource profiles is associated with differences in teachers' work-related outcomes. For example, in Collie et al.'s (2020a) study, the two Flourisher profiles (characterized by average or low demands, and high resources) tended to display the highest levels of job satisfaction (i.e., contentment regarding one's job) and occupational commitment (attachment to one's profession), followed by the two Average profiles (mixed or average resources/demands), and finally by the Struggler profile (high demands, low resources). At the school-level, Collie et al. (2020a) found that the Supportive school profile displayed higher levels on the two outcomes than the Unsupportive school profile. In addition, their results showed that Australian teachers and schools typically reported higher levels on these two outcomes than English teachers and schools. In the current study, we seek to replicate these findings. We also extended Collie et al.'s (2020a) results by considering the school profiles in relation to (school-average) student outcomes (perceptions of instructional support and achievement).

A growing body of research highlights the association between teachers' experiences at work and student outcomes. For example, teachers' job and personal resources have been associated with higher levels of student achievement, motivation, and perceived instructional support (Burić and Kim, 2020). The reverse has been found for job demands (Collie et al., 2020b). To our knowledge, however, prior research has almost universally employed variable-centered approaches to identify links between variables at a sample-wide level. One exception comes from Klusmann et al.'s (2008) research examining personal resource profiles (of engagement and resilience) among teachers. In this research, the authors revealed that more adaptive profiles (high engagement and resilience) tended to be associated with higher levels of perceived instructional support among students than less adaptive profiles.

In the present study, we rely on person-centered analyses to examine the extent to which the identified school demand-resource profiles will be associated with different levels of school-average instructional support as perceived by the students, and with different levels of school-average academic achievement in three areas (reading, mathematics, and science) as assessed in PISA 2018. We consider three types of instructional support: (1) Students' perceptions of autonomy-support, referring to the extent to which teachers are perceived as supporting students' empowerment and self-initiative (Skinner and Belmont, 1993; Ryan and Deci, 2017); (2) Students' perceptions of instrumental help, referring to the extent to which teachers are perceived as using feedback to guide student learning and improvement (Skinner and Belmont, 1993; Ang, 2005); (3) Students' perceptions of teacher warmth, referring to the extent to which teachers are perceived as displaying enjoyment and enthusiasm in relation to teaching and to the subject content (Skinner and Belmont, 1993; Keller et al., 2016). Together, the three types of instructional support encompass key components of well-known models of effective instruction (e.g., Skinner and Belmont, 1993; see also Hamre et al., 2013). Because school characteristics are known to be associated with students' perceptions of instructional support and achievement (e.g., Burns et al., 2019), we also examined the associations between the school profiles and student outcomes while controlling for school characteristics.

Overall, we anticipate that more supportive school profiles would be linked with higher levels on all outcomes. This is because schools where teachers experience more positive relationships and support, and where teachers are more confident in their teaching, are more likely to provide a more supportive and effective learning environment for students (e.g., Burić and Kim, 2020). Likewise, in schools where teachers feel supported at work, students are also more likely to feel supported and to thrive academically (Collie and Martin, 2017).

STUDY OVERVIEW

The aim of the present study was to replicate and extend prior research by identifying teacher and school demand-resource profiles, along with their predictors and outcomes at the teacher-level (Level 1) and school-level (Level 2). In the first (replication) phase of the study, demand-resource teacher and school profiles were identified among teachers from Australia and then England using multilevel latent profile analyses. Profile similarity tests were then conducted to systematically verify the extent to which the results are comparable across countries. Following this, additional analyses were conducted to assess whether and how the teacher and school profiles are predicted by teacher (sex, teaching experience) or school characteristics (school location, school size, proportion of students from low-SES backgrounds, and proportion of students from non-English speaking backgrounds; with and without controls for predictors), and are associated with teacher or school-average outcomes (job satisfaction and occupational commitment). In the second (extension) phase of the study relying on a subsample of the

Australian teachers, along with matched data from students, we ascertained the extent to which the school profiles are predicted by the same school characteristics and are associated with student outcomes (with and without controls for predictors in place). The student outcomes were students' perceptions of instructional support (i.e., autonomy-support, instrumental help, warmth) and achievement in reading, mathematics, and science.

METHODS

Sample and Procedure

Data used in the present study were from the TALIS 2018 (OECD, 2019c) and PISA 2018 (OECD, 2020) surveys. Institutional review board approval was received for the study and all ethical requirements were complied with in undertaking the study.

Phase 1

The Phase 1 sample included 5,439 teachers from 364 schools located in Australia and 2,216 teachers from 149 schools located in England from TALIS 2018. Every 5 years, the Organization for Economic Cooperation and Development (OECD) runs TALIS, which involves comprehensive and nationally-representative data collection among teachers in relation to a range of workplace experiences and perceptions. The TALIS 2018 sample was built using a two-stage probability sampling design to ensure a representative sample of schools and teachers in both countries (for details see OECD, 2019c).

Starting with the Australian sample, participating teachers were 62% female, had an average teaching experience of 15 (SD = 11) years, and around half (51%) of them were aged between 30 and 49 years. Most teachers (80%) were working full-time, and almost all (97%) had a bachelor's degree or higher. The Australian teachers taught at ISCED Level 2 (lower secondary) and/or ISCED Level 3 (upper secondary). Over half (61%) of the participating Australian schools were publicly managed, and the majority had fewer than one-third of students from low-SES backgrounds (67%). The majority of schools (57%) had a male principal who had, on average, 8 (SD = 7) years of experience as a principal. The schools were located in villages (6%; ≤3,000 people), towns (28%; 3,001–100,000 people), and cities (63%; >100,000 people). There were on average 15 (SD = 4) teachers per school.

Participating teachers from England were 65% female, had an average teaching experiences of 13 (SD = 12) years, almost two-thirds (62%) of them were aged between 30 and 49 years. Most teachers (80%) were working full-time, and 99% had a bachelor's degree or higher. The entire English sample taught at ISCED Level 2 (lower secondary). Just over one-third (36%) of the participating English schools were publicly managed, and the majority had fewer than one-third of students from low-SES backgrounds (69%). Most schools (59%) had a male principal who had, on average, 6 (SD = 5) years of experience as a principal. The schools were located in villages (9%; ≤3,000 people), towns (54%; 3,001–100,000 people), and cities (38%; >100,000 people). There were on average 15 (SD = 4) teachers per school.

Phase 2

A subsample of 2,099 Australian teachers from 130 schools was examined, along with matched data from 2,048 students who participated in PISA 2018. Using the OECD (2019c) TALIS-PISA link, it was possible to match the teacher and student data at the school level. The TALIS-PISA linked data were only available for the ISCED Level 3 (upper secondary) teachers in the Australian teacher sample (England did not participate in the TALIS-PISA link). Teachers from this subsample were 60% female, had an average teaching experience of 15 (SD = 11) years, and 51% of them were aged between 30 and 49 years. Most schools (59%) had a male principal who had, on average, 8 (SD = 7) years of experience as a principal. The schools were located in villages (8%; ≤3,000 people), towns (28%; 3,001–100,000 people), and cities (64%; >100,000 people). The students were 51% female with an average age of 15.80 years (SD = 0.30) years. In the subsample, there were on average 16 (SD = 4) teachers and 18 (SD = 5) students per school.

Measures

Teacher and school measures were drawn from the TALIS 2018 Teacher and Principal Questionnaires, respectively (OECD, 2019c). Student measures were drawn from the PISA 2018 Student Questionnaire (OECD, 2020; see **Supplementary Material** for items). The profile indicator variables and the teacher characteristics were modeled at the teacher-level. The teacher outcomes (job satisfaction and occupational commitment) were modeled at the teacher- and school-level. School characteristics and student outcomes were modeled only at the school-level.

Job Demands

Barriers to professional development was assessed with 6 items from the TALIS 2018 "Barriers to Professional Development" scale (e.g., "Professional development is too expensive/unaffordable"). *Disruptive student behavior* was assessed with items from the TALIS 2018 "Your Teaching" scale (3 items; e.g., "I lose quite a lot of time because of students interrupting the lesson"). For both scales, items were scored from 1 (Strongly disagree) to 4 (Strongly agree). Reliability was assessed with coefficient omega[1] and was adequate across both countries for barriers to professional development (ω = 0.755) and disruptive student behavior (ω = 0.908). The barriers to professional learning scale displayed 7% variance at the school-level (intraclass correlation [ICC] = 0.074). Although this is somewhat modest, this proportion is sufficient to support the need for multilevel analyses (Bliese et al., 2018). The disruptive student behavior scale also demonstrated adequate variance at the school-level (ICC = 0.169).

Job Resources

Teacher collaboration was assessed with items from the TALIS 2018 "Teaching in General" scale (3 items; e.g., "On average, how often do you do the following in this school? Exchange teaching materials with colleagues"). Items were scored on a scale

[1]Coefficient omega was calculated from the most invariant model in the measurement invariance tests and thus reflects the reliability across both countries.

from 1 (Never) to 6 (Once a week or more). Reliability estimates were adequate across both countries ($\omega = 0.719$) and there was adequate variance at the school-level (ICC = 0.195).

Teacher input in decision-making was assessed with items from the TALIS 2018 "School Climate" scale (3 items; e.g., "This school provides staff with opportunities to actively participate in school decisions"). Items were scored from 1 (Strongly disagree) to 4 (Strongly agree). Reliability was satisfactory ($\omega = 0.838$) and there was adequate variance at the school-level (ICC = 0.156).

Personal Resources

Teacher self-efficacy was assessed with items from the TALIS 2018 "Teaching in General" scale, which encompasses three types of self-efficacy: self-efficacy for classroom management (4 items; e.g., "Control disruptive behavior in the classroom"), self-efficacy for instruction (4 items; e.g., "Vary instructional strategies in my classroom"), and self-efficacy for student engagement (4 items; e.g., "Motivate students who show low interest in school work"). Items all followed the stem "In your teaching, to what extent can you do the following?" and were scored from 1 (Not at all) to 4 (A lot). Reliability for the three factors of self-efficacy was satisfactory ($\omega = 0.775$–855). For reasons of parsimony and because the self-efficacy factors were quite highly intercorrelated (r's = 0.60–0.63), self-efficacy was modeled as a single higher-order factor ($\omega = 0.824$) displaying modest, but sufficient variance at the school-level (ICC = 0.066; Bliese et al., 2018).

Teacher Characteristics

Teacher sex was coded 0 for female, 1 for male. *Teaching experience* was a continuous variable measured in years.

School Characteristics

School location was coded 1 for a village (≤3,000 people), 2 for a town (3,001–100,000 people), or 3 for a city (>100,000 people). School size was coded as 1 (under 250 students), 2 (250–400 students), 3 (500–749 students), 4 (750–999 students), or 5 (>1,000 students). Proportion of students from low socio-economic status (SES) backgrounds and proportion of students from non-English speaking backgrounds (NESB) were coded as 1 (0%), 2 (1–10%), 3 (11–30%), 4 (31–60%), or 5 (more than 60%).

Teacher Outcomes

Job satisfaction (3 items; e.g., "All in all, I am satisfied with my job") and *occupational commitment* (4 items; e.g., "If I could decide again, I would still choose to work as a teacher") were assessed with items from the TALIS 2018 "About Your Job" scale. For both outcomes, items were scored from 1 (Strongly disagree) to 4 (Strongly agree). Both outcomes were modeled at the teacher- and school-levels. Reliability was satisfactory at the teacher-level ($\omega = 0.847$ for job satisfaction; $\omega = 0.831$ for occupational commitment) and the school-level ($\omega = 0.989$ for job satisfaction; $\omega = 0.950$ for occupational commitment). Job satisfaction (ICC = 0.138) and occupational commitment (ICC = 0.047) both had enough variability at the school-level (Bliese et al., 2018).

Student Outcomes

Instructional support was assessed with three scales reported by students and aggregated at the school-level in analyses. *Autonomy-support* was assessed with the PISA 2018 "Teacher Support" scale (3 items; e.g., "The teacher listened to my view on how to do things"). Items were scored from 1 (Strongly disagree) to 4 (Strongly agree). Reliability was satisfactory at the student- ($\omega_{Student} = 0.904$) and school-level ($\omega_{School} = 0.990$), and school-level variability was sufficient (ICC = 0.036).

Instrumental help was assessed with items from the PISA 2018 "Perceived Feedback" scale (3 items; e.g., "The teacher gives me feedback on my strengths in this subject"). Items were scored from 1 (Never or almost never) to 4 (Every lesson or almost every lesson). Reliability was satisfactory at the student-level ($\omega_{Student} = 0.913$) and school-level ($\omega_{School} = 0.979$), and school-level variability was sufficient (ICC = 0.035).

Teacher warmth was assessed with the 4 items in the PISA 2018 "Perceived Teacher Interest" scale (4 items; e.g., "It was clear to me that the teacher liked teaching us"). Items were scored from 1 (Strongly disagree) to 4 (Strongly agree). Reliability was satisfactory at the student-level ($\omega_{Student} = 0.887$) and school-level ($\omega_{School} = 0.984$), and school-level variability was sufficient (ICC = 0.078).

School-average achievement in the areas of reading, mathematics, and science was measured *via* the Bayesian plausible values provided for each student in PISA 2018 and aggregated at the school-level in analyses. In PISA 2018, the reading test assessed students' capacity in locating information, comprehension and integrating knowledge, and evaluating and reflecting (OECD, 2019b). The mathematics test covered three areas: recognizing and identifying appropriate mathematical approaches; employing mathematical concepts and facts; and, interpreting, applying, and evaluating mathematical outcomes (for full details, see OECD, 2019b). The science test assessed three domains: understanding of different personal, local, national, and global contents; content, procedural, and epistemic knowledge relating to science facts, concepts, and theories; and, competencies including explaining, evaluating, and interpreting science tasks and data (OECD, 2019b). For each student and in each of the achievement areas, PISA produces 10 plausible values (for further details, see OECD, 2020). To accurately employ these scores, models need to be estimated 10 times (each with a different set of plausible values; OECD, 2020). Results are then aggregated using the Rubin (1987) strategy to obtain unbiased parameter estimates and standard errors. To calculate reliability, the 10 plausible values for each subject were used as indicators of a latent factor. Estimates were satisfactory at the student-level and school-level for reading achievement ($\omega_{Student} = 0.991$, $\omega_{Sschool} = 0.999$), mathematics achievement ($\omega_{Student} = 0.976$, $\omega_{Sschool} = 0.997$), and for science achievement ($\omega_{Student} = 0.984$, $\omega_{Sschool} = 0.998$). Reading achievement (ICC = 0.199), mathematics achievement (ICC = 0.245), and science achievement (ICC = 0.208) demonstrated adequate variance at the school-level.

Data Analysis

All analyses were conducted using M*plus* 8.4 (Muthén Muthén, 2019). Teacher (TCHWGT), student (W_FSTUWT), and school weights (SCHWGT) were applied to account for the probabilities of selection and participation at the different stages of sampling (see OECD, 2019b,c for details). The clustering of teachers within schools was accounted for in single-level modeling by using the M*plus* design-based correction procedures (Asparouhov, 2005). The robust maximum likelihood (MLR) estimator was used in all models. This estimator is robust to non-normality and to complex data structures. The limited amount of missing data was handled using full information maximum likelihood (FIML) estimation procedures (Enders, 2010). More precisely, missing data for teacher-level variables (demands, resources, teacher outcome, and teacher characteristics) were 1–5% (except disruptive student behavior, which was 16%). Missing data for school characteristics were 3–6%. Missing data for aggregated student-related outcomes were 7% for instructional support and 0% for achievement.

Preliminary Analyses

Preliminary confirmatory factor analyses (CFA) were conducted to ascertain the psychometric properties of our measures, and their measurement invariance (Millsap, 2011) across countries. Separate sets of models were estimated for the profile indicators (the five demands and resources), the teacher outcomes, and the students' instructional support outcomes. Because the profile indicator variables are only estimated at the teacher level, and then the profiles are used to estimate school-level profiles, the measurement models underpinning these indicators were estimated at the teacher level. In contrast, because the teacher outcomes were modeled at both the teacher and school levels, these preliminary analyses relied on multilevel-CFA. Finally, because student outcomes were reported by the students, but used at the school level (student and teachers were only matched at the school level in the TALIS-PISA link), these preliminary analyses relied on multilevel-CFA conducted at the student and school levels. The estimates of composite reliability (omega) and intraclass correlation coefficients (ICC) reported in the "Measures" sections were calculated from the most invariant of these measurement models (factor loadings, intercepts, residuals, covariances, variances, and means). Factor scores were saved from the most invariant of these measurement models and used as input for our main analyses, together with manifest scores reflecting the teacher characteristics, school characteristics, and aggregated achievement outcomes. Additional details on these preliminary analyses, which supported the complete invariance of our measures and the isomorphism (equality) of our factor loadings across levels for all multilevel analyses are reported in the first section of the **Supplementary Material**. **Table 1** shows the reliability coefficients and descriptive statistics at the teacher-level and school-level. Latent correlations among the variables from the most invariant models are available in **Supplementary Table 2**. Prior to undertaking our main analyses, a multigroup (across countries) baseline model using the factor scores of the profile indicators was estimated to standardize the sampling weights separately for each country using procedures outlined in Collie et al. (2020a). These standardized weights were then used in all analyses outlined below.

Phase 1: Single-Level and Multilevel LPA

This initial phase of the study seeks to replicate Collie et al.'s (2020a) study using analytical procedures extensively documented in that previous study. For this reason, we only provide a brief summary of the analytic steps here, and refer interested readers to the **Supplementary Material** of the Collie et al. (2020a) study for additional details. First, single-level latent profile analyses (LPA) were used to estimate teacher profiles based on the means and variance of the profile indicators at Level 1 (L1). Once the optimal representation of teachers' profile was identified, multilevel-LPA were used to identify school profiles based on the relative frequency of occurrence of these teacher profiles at the school level. For both single-level and multilevel LPA, we estimated solutions including 1 to 8 profiles, separately for the two countries. Each model was estimated using 10,000 random sets of start values, 1,000 iterations, and 100 final stage optimizations. We verified that the best log-likelihood value was properly replicated for all models.

Several indices were employed to assess the relative adequacy of the models, along with elbow plots reflecting the decrease in the value of these indicators as a function of added profiles (e.g., Morin and Litalien, 2019). More precisely, we relied on the Akaike information criterion (AIC) and its consistent version (CAIC), and on the Bayesian information criterion (BIC) and its sample size-adjusted version (SSA-BIC). For these indices, a lower value indicates better fit. Because these values are sample size dependent and thus often fail to converge on a specific solution, we consider a graphical display of these indicators (i.e., an elbow plot). The point at which the drop in the value of these indicators noticeably flattens can be used to guide model selection (Morin et al., 2016). For the single-level models, we also report the *p*-value associated with the adjusted Lo–Mendel–Rubin Likelihood Ratio Test (*p*LMR; this indicator is not available for multilevel-LPA). A statistically significant value on this test supports the value of a solution in relation to a solution with one fewer profile. Alongside these indices, we used parsimony, conceptual relevance, and statistical adequacy to help determine the optimal solution.

After determining the optimal solution separately for each country, at both levels, we undertook tests of profile similarity to ascertain the extent to which the profile solutions could be considered to be comparable across the two countries (Morin et al., 2016). These tests were first conducted for the single-level LPA (Morin et al., 2016) and then for the multilevel LPA starting from the most similar single-level LPA solution (Collie et al., 2020a). More precisely, at the teacher level, we estimated models of configural (i.e., same numbers of profiles), structural (i.e., same within-profile means on the profile indicators), dispersion (i.e., same within-profile means and variances on the profile indicators), and distributional (i.e., same within-profile means and variances, and same profile sizes) similarity across countries (Morin et al., 2016). At the school level, we estimated models of configural, structural, and distributional similarity (Collie et al., 2020a), as no variance components are involved in the estimation

TABLE 1 | Reliabilities and descriptive statistics for both countries.

		Australia		England	
	ω	*M*	*SD*	*M*	*SD*
Teacher-level					
Barriers to professional development	0.755	2.247	0.579	2.349	0.583
Disruptive student behavior	0.908	2.092	0.775	2.011	0.776
Teacher collaboration	0.719	4.962	0.971	4.662	1.063
Teacher input	0.838	2.755	0.620	2.719	0.595
Teacher self-efficacy	0.824	3.246	0.466	3.340	0.451
Job satisfaction	0.847	3.157	0.589	3.022	0.612
Occ. commitment	0.831	3.122	0.628	2.841	0.691
School-level					
Job satisfaction	0.989	3.153	0.253	3.020	0.288
Occ. commitment	0.950	3.126	0.210	2.844	0.230
Autonomy-support	0.990	2.865	0.239	—	—
Instrumental help	0.979	2.600	0.263	—	—
Teacher Warmth	0.984	1.786	0.286	—	—
Reading achievement	0.999	497	53	—	—
Mathematics achievement	0.997	486	44	—	—
Science achievement	0.998	498	47	—	—

For all teacher-related factors, coefficient omega (ω) is reported for both countries as it was taken from the most invariant measurement invariance test. For the student outcomes, omega involves only the Australian sample for the TALIS-PISA link (given England did not participate in the TALIS-PISA link; however, see Sizmur et al., 2019 for details about school-level outcomes for England). The achievement values are reported as whole numbers as per OECD guidelines (OECD, 2019b). Occ. commitment, Occupational commitment.

of L2 profiles defined on the basis of the relative frequency of occurrence of L1 profiles.

Three additional tests were then estimated to examine the equivalence of the associations between: (a) the predictors (i.e., teacher characteristics at L1: sex and teaching experience; school characteristics at L2: school location, school size, proportion of low-SES students, and proportion of NESB students) and the likelihood of profile membership (predictive similarity; Morin et al., 2016); (b) the profiles and the outcomes (i.e., L1 and L2 job satisfaction and occupational commitment; explanatory similarity; Morin et al., 2016); and (c) the profiles and the outcomes while controlling for the effects of the predictors (adjusted-explanatory similarity; derived from the profile-based ANCOVA approach of McLarnon and O'Neill, 2018). The L2 predictive similarity tests and the L1 and L2 adjusted-explanatory similarity tests represent extensions to the Collie et al. (2020a) study. We provide annotated input syntax for these tests in the **Supplementary Material**.

For these tests, profile similarity is supported as long as two out of the four aforementioned information criteria are lower for a solution when compared to the previous one from the sequence of profile similarity tests. Predictors and outcomes were directly incorporated into the retained (most similar) solution from the previous steps at L1 as per Morin et al. (2016), and at L2 using the manual three-step approach developed by Collie et al. (2020a) from early work conducted by Litalien et al. (2019) and Morin and Litalien (2017). This approach at L2 was necessary given the way the multilevel LPA was conducted (where the L1 profiles are "predicted" by the L2 profiles; see Collie et al., 2020a). Associations with predictors were assessed using a multinomial logistic link function (i.e., the impact of predictors on the likelihood of membership into each profile relative to all other profiles was estimated; Vermunt, 2010), whereas associations with outcomes were assessed using mean comparisons implemented with the multivariate delta method (Raykov and Marcoulides, 2004). Associations with outcomes while controlling for predictors used the same methods, but also involved regressing the outcomes on the background characteristics (McLarnon and O'Neill, 2018). For interested readers, annotated input files are provided in the Supplementary Material of Collie et al.'s (2020a) article, except for tests of L2 predictive similarity and L1 and L2 adjusted-explanatory similarity, which are provided in our **Supplementary Material**.

Phase 2: School-Level Profiles and Student Outcomes in the Australian Subsample

Phase 2 involved an extension of the Collie et al. (2020a) study to consider how the school profiles identified in Phase 1 were associated with predictors and student outcomes among a subsample of the Australian teachers. The baseline model used in these analyses was specified (i.e., constrained) to be identical to the most similar model retained in Phase 1, using the manual three-step procedures outlined in Collie et al. (2020a). Then, three separate analyses were conducted that paralleled Phase 1. For the first analysis, we examined the extent to which the four school characteristics (i.e., school location, school size, proportion of students from low-SES backgrounds, and proportion of students from NESB backgrounds) predicted the likelihood of profile membership (Vermunt, 2010).

For the second analysis, we examined the extent to which the school-level profiles are associated with different levels of the student-reported outcomes using mean comparisons

implemented with the multivariate delta method (Raykov and Marcoulides, 2004). The model was run 10 times to account for the 10 Bayesian plausible values provided in PISA for each of the achievement outcomes. The estimates and *p*-values were pooled using the Rubin (1987) strategy (which is automated in the M*plus* 8.4 statistical package; Muthén Muthén, 2019) to obtain unbiased parameter estimates. The third analysis examined associations between profile membership and the student outcomes while controlling for the school characteristics using the previously described approach (McLarnon and O'Neill, 2018).

RESULTS

Phase 1: Single-Level LPA

Profile Identification and Description

Model fit statistics for the solutions involving 1 through 8 profiles estimated separately in Australia and England are reported in **Table 2**, and corresponding elbow plots are reported in **Supplementary Figure 1**. For both countries, all information criteria kept on decreasing with the addition of profiles to the solution. The *p*LMR suggested a 6 profile solution in Australia, and either a 4, or 6 profile solution in England. Examination of the elbow plots revealed a slight flattening around 5-and 6 profiles in both countries. In summary, these statistics generally suggest that the optimal solution should include somewhere between 5 and 6 profiles. These two possible solutions, together with the adjacent 4 and 7 profile solutions, were thus further examined for conceptual relevance, parsimony, and meaningfulness. Examination of the 6 profile solution revealed a high level of similarity across country, thus providing preliminary support for configural similarity. When we compared the 5 profile solution to the 6 profile solution, it was clear that the additional profile was meaningful in both countries and presented a differentiated shape compared with the other profiles. In contrast, the 4 profile solution lacked the nuance that was evident in the other solutions, whereas the 7 profile solution did not add anything new (simply resulting in the arbitrary division of one profile into two very similar ones). The 6 profile solution was thus retained for both countries, and submitted to more systematic tests of profile similarity. The results from the L1 tests of profile similarity conducted across the two countries are reported in **Table 3**, and support the complete similarity (configural, structural, dispersion, and distributional) of the solution across Australia and England. A graphical representation of this final 6 profile solution of distributional similarity is presented in **Figure 2**, and detailed results associated with this solution are reported in **Supplementary Table 6**.

Teachers corresponding to Profile 1 (11% of the sample) reported low barriers to professional development, very low disruptive behavior, above average teacher collaboration and teacher input, and high self-efficacy. This profile was thus labeled *Low-Demand-Flourisher* to reflect this adaptive blend of low job demands and high job and personal resources. Teachers corresponding to Profile 2 (17% of the sample) reported very low barriers to professional development, below average disruptive behavior, and high teacher collaboration, teacher input, and self-efficacy. This profile was labeled *Mixed-Demand-Flourisher* because of the consistently high levels of the job and personal resources.

Teachers corresponding to Profile 3 (11% of the sample) reported low barriers to professional development, slightly below average disruptive behavior, high teacher collaboration, high teacher input, and above average self-efficacy. This profile was thus labeled *Job-Resourced-Average* to reflect the low to average job demands, coupled with average self-efficacy, but high job resources. Teachers corresponding to Profile 4 (14% of the sample) reported slightly above average barriers to professional development, slightly below average disruptive behavior, average teacher collaboration, and slightly below average teacher input and self-efficacy. We labeled this profile *Balanced-Average* to reflect the matching average levels observed across all demands and resources.

Teachers corresponding to Profile 5 (11% of the sample) reported high barriers to professional development and disruptive behavior, slightly below average teacher collaboration, high teacher input, and low self-efficacy. We labeled this profile *Mixed-Resourced-Struggler* to reflect the mixed levels of resources and high demands. Teachers corresponding to Profile 6 (36% of the sample) reported high barriers to professional development and disruptive behavior, and low teacher collaboration, teacher input, and self-efficacy. We labeled this profile *Low-Resourced-Struggler* to reflect this blend of high job demands, and low job and personal resources.

Profile Prediction and Outcomes

The results from the tests of predictive (associations with predictors), explanatory (associations with outcomes), and adjusted-explanatory (associations with outcomes controlling for predictors) similarity (**Table 3**) supported the equivalence of the associations between the profiles, their predictors, and their outcomes both without and with controls for the predictors in place. The results from these analyses are reported in **Table 4** for the predictors, and in **Table 5** for the outcomes and the outcomes adjusted for the predictors.

For the predictors, male teachers were less likely to correspond to the two types of *Flourisher* profiles and to the *Job-Resourced-Average* profile than to the *Balanced-Average* or *Low-Resourced-Struggler* profiles. Teachers with greater teaching experience were more likely to correspond to the *Low-Demand-Flourisher* profile than all other profiles. Teachers with greater teaching experience were also more likely to correspond to the *Mixed-Demand-Flourisher*, *Job-Resourced-Average*, and the *Balanced-Average* profiles than to the two types of *Struggler* profiles. Taken together, these results suggest that male teachers and less experienced teachers were more likely to be in the arguably less desirable profiles.

For the outcomes, members of the *Mixed-Demand-Flourisher* profile displayed the highest levels of job satisfaction and occupational commitment, followed by members of the *Low-Demand-Flourisher* profile. The next highest level was observed in the *Job-Resourced-Average* profile, followed equally by the *Balanced-Average* profile and *Mixed-Resourced-Struggler* profile. Finally, the *Low-Resourced-Struggler* profile displayed

TABLE 2 | Fit statistics and entropy for Australia and England.

	Log-likelihood	Free Parameters	AIC	CAIC	BIC	SSA-BIC	*p*LMR	Entropy
Australia—Single-level								
1 profile	−35132.680	10	70285.361	70361.374	70351.374	70319.597	—	—
2 profiles	−33104.629	21	66251.259	66410.887	66389.887	66323.156	<0.01	0.845
3 profiles	−31391.072	32	62846.143	63089.387	63057.387	62955.701	<0.01	0.771
4 profiles	−30259.995	43	60605.990	60932.848	60889.848	60753.208	<0.01	0.810
5 profiles	−29249.431	54	58606.862	59017.335	58963.335	58791.740	<0.01	0.823
6 profiles	−28773.371	65	57676.743	58170.830	58105.830	57899.281	<0.01	0.832
7 profiles	−28388.420	76	56928.841	57506.543	57430.543	57189.039	<0.01	0.843
8 profiles	−28121.379	87	56416.757	57078.075	56991.075	56714.616	*ns*	0.826
Australia—Multilevel								
1 profile	−9251.758	5	18513.517	18551.523	18546.523	18530.635	—	0.709
2 profiles	−9174.344	11	18370.688	18454.303	18443.303	18408.349	—	0.681
3 profiles	−9127.009	17	18288.018	18417.241	18400.241	18346.221	—	0.670
4 profiles	−9113.675	23	18273.349	18448.180	18425.180	18352.093	—	0.696
5 profiles	−9102.291	29	18262.580	18483.019	18454.019	18361.867	—	0.676
6 profiles	−9093.576	35	18257.152	18523.199	18488.199	18376.980	—	0.692
7 profiles	−9085.710	41	18253.421	18565.076	18524.076	18393.791	—	0.693
8 profiles	−9078.078	47	18250.155	18607.419	18560.419	18411.068	—	0.702
England—Single-level								
1 profile	−14418.504	10	28857.008	28924.042	28914.042	28882.271	—	—
2 profiles	−13573.799	21	27189.598	27330.371	27309.371	27242.651	<0.01	0.838
3 profiles	−12846.487	32	25756.973	25971.484	25939.484	25837.815	<0.01	0.890
4 profiles	−12330.889	43	24747.779	25036.028	24993.028	24856.410	<0.01	0.809
5 profiles	−11988.036	54	24084.071	24446.058	24392.058	24220.492	*ns*	0.821
6 profiles	−11801.073	65	23732.147	24167.871	24102.871	23896.357	*ns*	0.825
7 profiles	−11660.384	76	23472.767	23982.230	23906.230	23664.767	<0.01	0.819
8 profiles	−11484.460	87	23142.920	23726.121	23639.121	23362.709	*ns*	0.830
England—Multilevel								
1 profile	−3712.827	5	7435.654	7469.171	7464.171	7448.286	—	0.720
2 profiles	−3682.214	11	7386.428	7460.166	7449.166	7414.217	—	0.697
3 profiles	−3673.464	17	7380.929	7494.887	7477.887	7423.876	—	0.667
4 profiles	−3668.075	23	7382.150	7536.330	7513.330	7440.255	—	0.676
5 profiles	−3663.898	29	7385.796	7580.197	7551.197	7459.059	—	0.681
6 profiles	−3660.468	35	7390.935	7625.556	7590.556	7479.356	—	0.696
7 profiles	−3658.564	41	7399.127	7673.969	7632.969	7502.705	—	0.707
8 profiles	−3657.222	47	7408.445	7723.507	7676.507	7527.181	—	0.693

AIC, Akaike Information Criteria; CAIC, Consistent Akaike Information Criteria; BIC, Bayesian Information Criteria; SSA-BIC, sample-size-adjusted Bayesian Information Criteria; pLMR, p-value of the Lo–Mendell–Rubin Likelihood Ratio Test; ns, non-significant.

the lowest levels on both outcomes. For the outcomes after correcting for the predictors, the pattern of results was the same, with one exception: the *Mixed-Demand-Flourisher* and the *Low-Demand-Flourisher* exhibited similar levels of occupational commitment after controlling for teachers' background characteristics.

Phase 1: Multilevel LPA

Profile Identification and Description

Model fit statistics for the multilevel solutions with 1 through 8 profiles estimated separately in Australia and England are reported in the bottom section of **Table 2** and corresponding elbow plots are reported in **Supplementary Figure 2**. For Australia, the 3 profile L2 solution resulted in the lowest value for the CAIC, BIC, and SSA-BIC, whereas the AIC continued to decrease up to 8 profiles. However, the elbow plot showed a flattening in the decrease of the value of the AIC and SSA-BIC around three profiles, although this flattening was first apparent in relation to the 2 profile solution for the CAIC and BIC. For England, the 3 profile L2 solution resulted in the lowest value of the AIC, whereas the 2 profile solution resulted in the lowest values for the CAIC, BIC, and SSA-BIC. For the AIC, the elbow plot showed a flattening at two profiles. In summary, these statistics generally suggested than

TABLE 3 | Tests of profile similarity across Australia and England.

	Log-Likelihood	Free Parameters	AIC	CAIC	BIC	SSA-BIC	Entropy
Single-level LPA							
Configural	−45218.140	131	90698.281	91738.828	91607.829	91191.538	0.876
Structural (means)	−45336.164	101	90874.328	91676.583	91575.583	91254.626	0.878
Dispersion (means and variances)	−45367.694	71	90877.387	91441.349	91370.348	91144.725	0.878
Distributional (means, variances, size)	−45388.761	66	90909.523	91433.768	91367.768	91158.034	0.877
Predictive similarity							
Unconstrained across country	−45027.831	26	90107.662	90314.054	90288.054	90205.431	0.878
Constrained across country	−45047.773	16	90127.546	90254.556	90238.556	90187.711	0.878
Explanatory similarity							
Unconstrained across country	−53597.303	32	107258.606	107512.786	107480.785	107379.096	0.888
Constrained across country	−53635.126	20	107310.253	107469.114	107449.115	107385.559	0.888
Adjusted-explanatory similarity							
Unconstrained across country	−50513.698	47	101121.395	101494.488	101447.488	101298.132	0.885
Constrained across country	−50553.427	35	101176.854	101454.689	101419.689	101308.466	0.885
Multilevel LPA							
Configural	−13172.195	23	26390.390	26573.082	26550.082	26476.993	0.754
Structural (proportion of L1 profiles)	−13191.688	13	26409.376	26512.636	26499.637	26458.325	0.745
Distributional (proportion of L2 profiles)	−13194.335	12	26412.670	26507.987	26495.987	26457.854	0.744
Predictive similarity							
Unconstrained across country	−16218.676	18	32473.353	32616.328	32598.329	32541.129	0.767
Constrained across country	−16236.084	14	32500.169	32611.372	32597.373	32552.883	0.757
Explanatory similarity							
Unconstrained across country	−12434.810	12	24893.621	24988.937	24976.938	24938.805	0.792
Constrained across country	−12545.187	8	25106.375	25169.919	25161.919	25136.497	0.785
Adjusted-explanatory similarity							
Unconstrained across country	−15263.750	33	30593.501	30717.756	30822.624	30855.623	0.771
Constrained across country	−15366.656	29	30791.313	30900.507	30992.663	31021.662	0.773

AIC, Akaike Information Criteria; CAIC, Consistent Akaike Information Criteria; BIC, Bayesian Information Criteria; SSA-BIC, sample-size-adjusted Bayesian Information Criteria.

the optimal solution should have 2 or 3 profiles. These two solutions were thus further examined for conceptual relevance, parsimony, and meaningfulness. This examination revealed that the profile added as part of the 3 profile solutions did not meaningfully differ from the profiles already included in the 2 profile solution in either country. Accordingly, a solution comprising 2 school-level profiles was retained in both countries.

The results from the L2 tests of profile similarity conducted across the two countries are reported in the bottom section of **Table 3**, and support the complete similarity (configural, structural, and distributional) of the solution across Australia and England. A graphical representation of the final 2 profile solution of L2-distributional similarity is presented in **Figure 3**. Examination of this solution suggested the presence of an Unsupportive school profile (43% of the schools) and a Supportive school profile (57% of the schools). The *Unsupportive school profile* included a high proportion of members from the *Low-Resourced-Struggler* (48%) profile, followed by the *Mixed-Resourced-Struggler* (14%), *Balanced-Average* (12%), *Mixed-Demand-Flourisher* (11%), *Job-Resourced-Average* (9%), and *Low-Demand-Flourisher* (6%) profiles. Turning to the second school profile, the *Supportive school profile* included a high proportion of the *Mixed-Demand-Flourisher* (24%), followed by the *Low-Resourced-Struggler* (23%), the *Low-Demand-Flourisher* and *Balanced-Average* profiles (both 17%), the *Job-Resourced-Average* (11%), and finally the *Mixed-Resourced-Struggler* (8%).

Profile Prediction and Outcomes

The results from the tests of predictive (associations with predictors), explanatory (associations with outcomes), and adjusted-explanatory (associations with outcomes controlling for predictors) similarity are shown in **Table 3**. These results support the equivalence of the associations between the profiles and their predictors across countries, but revealed differences across countries related to the associations between the profiles and the outcomes (with and without controls). The results from these analyses are reported in **Table 6** for the predictors, and in **Table 7** for the outcomes and adjusted outcomes. For the predictors, the results reveal a single statistically significant association suggesting that schools with a higher proportion of low-SES students were more likely to correspond to the *Unsupportive* school profile than to the *Supportive* one.

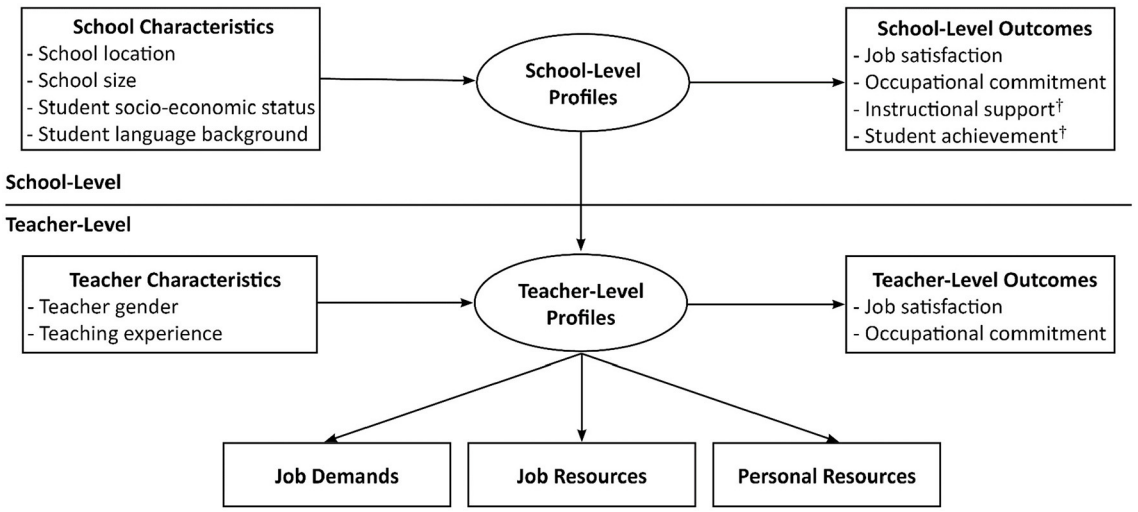


FIGURE 1 | Hypothesized models tested in the study at the teacher- and school-level. In phase one of analysis, teacher-level profiles were identified based on the demands and resources. Then, tests of associations between profile membership and predictors (teacher characteristics) and outcomes (teacher-level outcomes) were conducted. Following this, school-level profiles were identified. Then, tests of associations between profile membership and predictors (school characteristics) and outcomes (school-level outcomes) were conducted. In phase two†, analyses involved a subsample of Australian teachers who worked at schools with matched student data, and tests between profile membership and student outcomes (i.e., instructional support, student achievement) were conducted. Not shown here are the tests of profile similarity that were conducted to compare the teacher- and school-level results across countries (see Methods).

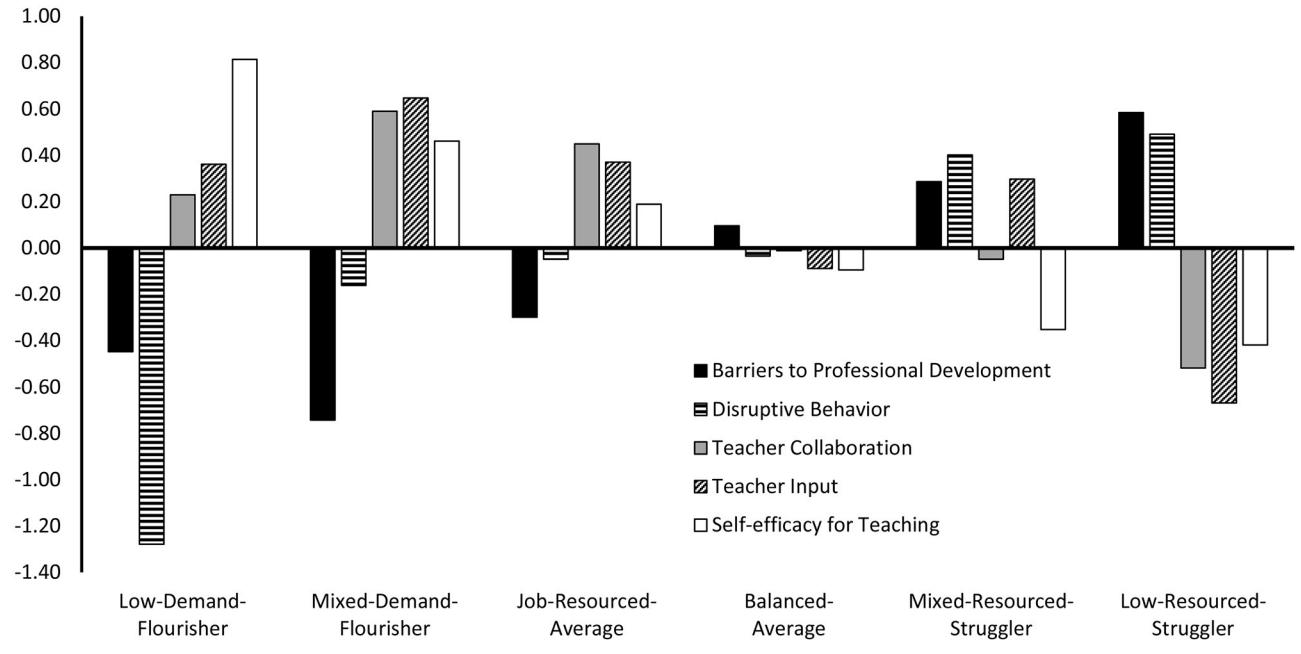


FIGURE 2 | Single-level LPA results of distributional similarity showing teacher profiles for both countries.

For the outcomes, due to the lack of explanatory similarity across country, we compared the school-level means of the teacher outcomes within and across Australia and England. In both countries, the *Supportive* school profile displayed significantly higher school-average job satisfaction and occupational commitment than the *Unsupportive* one. When comparing matched profiles across the two countries, the two school profiles from the Australian sample displayed significantly higher levels on the two outcomes than the matching profiles estimated in the English sample. For the outcomes adjusted for school characteristics, the pattern of results was the same as that found for the unadjusted outcomes.

TABLE 4 | The role of teacher characteristics in predicting profile membership in both countries (Single-level LPA).

	b	*SE*	*OR*	*b*	*SE*	*OR*
	Low-Demand-Flourisher vs. Mixed-Demand-Flourisher			**Low-Demand-Flourisher vs. Job-Resourced-Average**		
Sex (F/M)	−0.042	0.107	0.959	−0.073	0.121	0.929
Teaching experience	0.039**	0.005	1.040	0.039**	0.005	1.039
	Low-Demand-Flourisher vs. Balanced-Average			**Low-Demand-Flourisher vs. Mixed-Resourced-Struggler**		
Sex (F/M)	−0.442**	0.107	0.643	−0.237	0.121	0.789
Teaching experience	0.038**	0.005	1.039	0.057**	0.005	1.058
	Low-Demand-Flourisher vs. Low-Resourced-Struggler			**Mixed-Demand-Flourisher vs. Job-Resourced-Average**		
Sex (F/M)	−0.314**	0.197	0.730	−0.032	0.116	0.979
Teaching experience	0.049**	0.004	1.050	−0.001	0.005	0.999
	Mixed-Demand-Flourisher vs. Balanced-Average			**Mixed-Demand-Flourisher vs. Mixed-Resourced-Struggler**		
Sex (F/M)	−0.400**	0.102	0.670	−0.195	0.112	0.823
Teaching experience	−0.001	0.005	0.999	0.017**	0.005	1.017
	Mixed-Demand-Flourisher vs. Low-Resourced-Struggler			**Job-Resourced-Average vs. Balanced-Average**		
Sex (F/M)	−0.273**	0.087	0.761	−0.369**	0.111	0.692
Teaching experience	0.009*	0.004	1.009	0.001	0.006	1.000
	Job-Resourced-Average vs. Mixed-Resourced-Struggler			**Job-Resourced-Average vs. Low-Resourced-Struggler**		
Sex (F/M)	−0.163	0.128	0.849	−0.241*	0.095	0.786
Teaching experience	0.018**	0.006	1.018	0.010*	0.004	1.010
	Balanced-Average vs. Mixed-Resourced-Struggler			**Balanced-Average vs. Low-Resourced-Struggler**		
Sex (F/M)	0.205	0.119	1.228	0.128	0.088	1.136
Teaching experience	0.018**	0.006	1.018	0.010*	0.004	1.010
	Mixed-Resourced-Struggler vs. Low-Resourced-Struggler					
Sex (F/M)	−0.078	0.104	0.925			
Teaching experience	−0.008	0.005	0.992			

*$p \leq 0.05$; ** $p \leq 0.01$; *b*, multinomial logistic regression coefficient; *SE*, standard error of the coefficient; *OR*, odds ratio. For sex, females were coded 0 and males were coded 1.

TABLE 5 | Means of teacher-level outcomes from explanatory similarity test and adjusted-explanatory similarity test (Single-level LPA).

	Low-Demand—Flourisher *M* (95% CI)	**Mixed-Demand—Flourisher *M* (95% CI)**	**Job-Resourced—Average *M* (95% CI)**	**Balanced—Average *M* (95% CI)**	**Mixed-Resourced—Struggler *M* (95% CI)**	**Low-Resourced—Struggler *M* (95% CI)**
Explanatory similarity test						
Job satisfaction	0.217 (0.181, 0.253)	0.321 (0.296, 0.347)	0.096 (0.062, 0.129)	−0.040[a] (−0.071, −0.008)	−0.052[a] (−0.087, −0.017)	−0.263 (−0.283, −0.243)
Occupational commitment	0.250 (0.216, 0.284)	0.358 (0.334, 0.381)	0.122 (0.091, 0.154)	−0.023[a] (−0.052, 0.006)	0.003[a] (−0.027, 0.032)	−0.307 (−0.325, −0.288)
Adjusted-explanatory similarity test (with controls for predictors)						
Job satisfaction	0.267 (0.219, 0.315)	0.308 (0.265, 0.351)	0.159 (0.117, 0.202)	0.019[a] (−0.024, 0.062)	0.003[a] (−0.040, 0.046)	−0.146 (−0.184, −0.107)
Occupational commitment	0.246[b] (0.200, 0.292)	0.306[b] (0.268, 0.345)	0.127 (0.087, 0.166)	−0.026[a] (−0.067, 0.015)	0.005[a] (−0.036, 0.045)	−0.269 (−0.303, −0.236)

Superscript values indicate a mean comparison that was not significantly different across the profiles with the same superscript value across a row. All other comparisons were significant at $p < 0.05$.

Phase 2: School Profiles and Instructional Support and School-Average Achievement

Phase 2 involved examining the extent to which the school-level profiles are associated with student outcomes (i.e., perceived instructional support and academic achievement). An initial model was estimated including the four school characteristics as predictors of school profile membership. As before, the results from the multinomial logistic regressions revealed only one

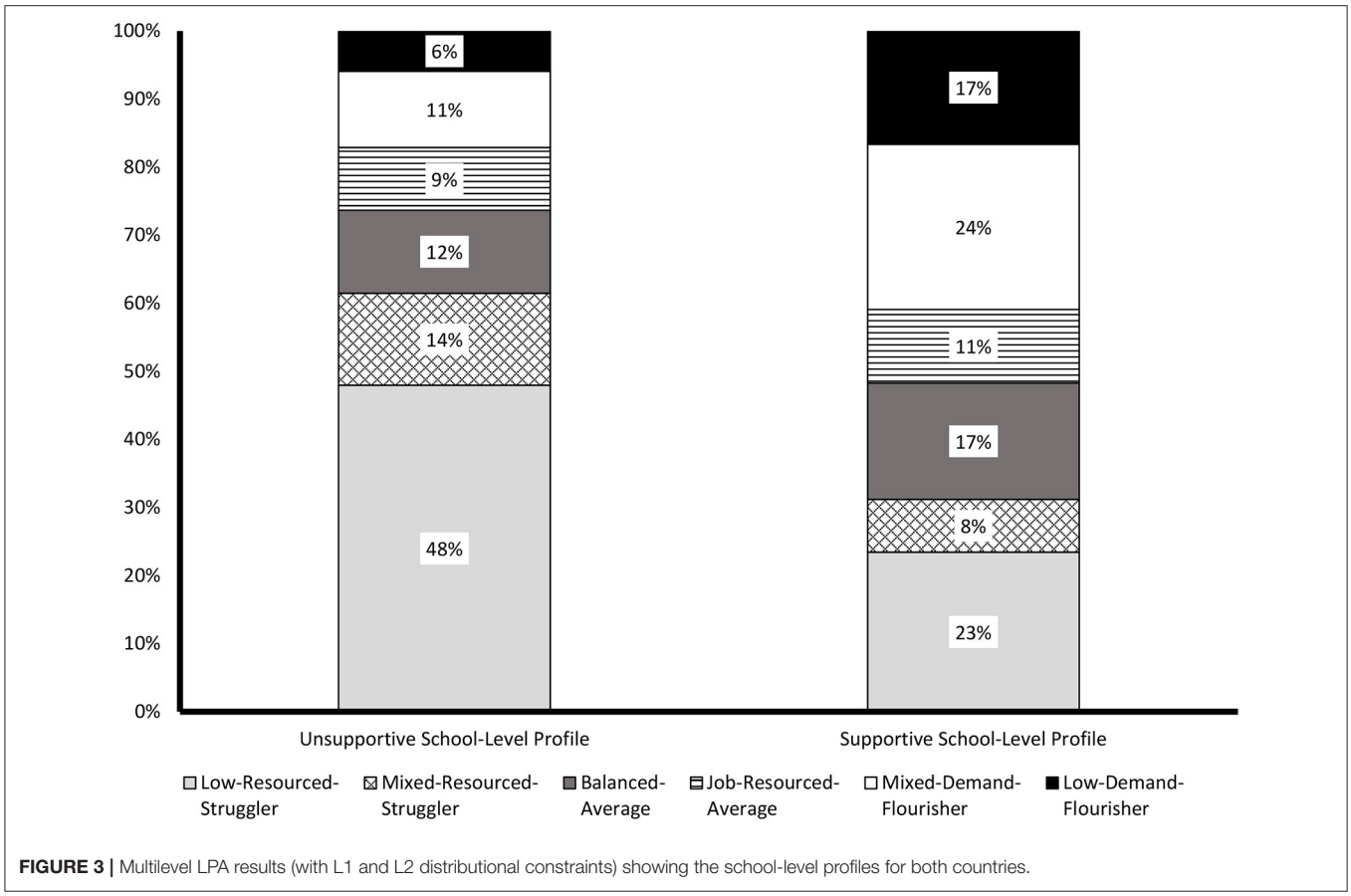


FIGURE 3 | Multilevel LPA results (with L1 and L2 distributional constraints) showing the school-level profiles for both countries.

TABLE 6 | The role of school characteristics in predicting membership in the supportive school profile in both countries (Multilevel LPA).

	b	*SE*	*OR*
School location	0.510	0.475	1.665
School size	0.070	0.186	1.073
Proportion students from low-SES backgrounds	−1.054**	0.249	0.349
Proportion students from NESB backgrounds	0.380	0.272	1.462

***p ≤ 0.01; b, multinomial logistic regression coefficient; SE, standard error of the coefficient; OR, odds ratio; SES, socio-economic-status; NESB, non-English speaking backgrounds. For school location and school size, higher values represent more populous locations and larger schools, respectively. For proportion of students from low-SES or NESB backgrounds, larger values represent higher proportions in a school.*

statistically significant association, showing that schools with a higher proportion of low-SES students were more likely to correspond to the *Unsupportive* profile than to the *Supportive* one ($b = 1.69$, $SE = 0.74$, $p = 0.02$, odds ratio $= 5.44$).

Next, the associations among the profiles and the school-average outcomes were examined. These results are presented in **Table 8** and first reveal that the *Supportive* school profile was characterized by significantly higher levels of school-average autonomy-support, instrumental help, and teacher warmth when compared to the *Unsupportive* school profile. The *Supportive* school profile was also characterized by significantly higher school-average levels of reading, mathematics, and science achievement than the *Unsupportive* school profile. Finally, these associations were re-estimated while controlling for the school characteristics (**Table 8**). The results from these analyses where identical to those previously discussed without including this control.

DISCUSSION

Our study was designed to replicate and extend Collie et al.'s (2020a) research. In Phase 1 (the replication phase), many of the findings obtained by Collie et al. (2020a) were reproduced. In particular, our results led to the identification of the same set of school profiles (a Supportive school profile and an Unsupportive school profile), and revealed a generally similar pattern of associations with predictors and outcomes at the teacher and school level. However, one notable difference was the identification of six, rather than five, teacher profiles: Low-Demand-Flourisher, Mixed-Demand-Flourisher, Job-Resourced-Average, Balanced-Average, Mixed-Resourced-Struggler, and Low-Resourced-Struggler. More precisely, the current results led to the identification of five profiles matching those identified by Collie et al. (2020a), along with the identification of one additional Struggler profile (Mixed-Resourced-Struggler)

TABLE 7 | Means of school-level outcomes from explanatory similarity test and adjusted-explanatory similarity test (Multilevel LPA).

	Unsupportive School Profile *M* (95% CI)		Supportive School Profile *M* (95% CI)	
	Australia	**England**	**Australia**	**England**
Explanatory similarity test				
School-average job satisfaction	−0.114 (−0.158, −0.070)	−0.233 (−0.276, −0.190)	0.148 (0.094, 0.203)	0.055 (0.028, 0.083)
School-average occupational commitment	−0.039 (−0.062, −0.016)	−0.203 (−0.233, −0.173)	0.109 (0.068, 0.150)	−0.027 (−0.047, −0.007)
Adjusted-explanatory similarity test (with controls for predictors)				
School-average job satisfaction	−0.055 (−0.154, 0.044)	−0.162 (−0.268, −0.057)	0.140 (0.062, 0.217)	0.075 (−0.010, 0.160)
School-average occupational commitment	0.014 (−0.043, 0.072)	−0.135 (−0.202, −0.068)	0.117 (0.063, 0.172)	−0.007 (−0.062, 0.048)

All comparisons within and across country were significantly different at $p < 0.05$ for both tests.

TABLE 8 | Means of student-reported outcomes related to instructional practice and achievement (Phase 2).

	Unsupportive school profile *M* (95% CI)	Supportive school profile *M* (95% CI)	*p*-value
Model (with no covariate controls)			
Instructional support			
Autonomy-support	−0.052 (−0.078, −0.026)	0.043 (0.019, 0.068)	<0.01
Instrumental help	−0.032 (−0.056, −0.008)	0.027 (0.004, 0.050)	<0.01
Teacher warmth	−0.077 (−0.115, −0.038)	0.064 (0.027, 0.100)	<0.01
School-average achievement			
Reading achievement	456 (443, 470)	531 (519, 544)	<0.01
Mathematics achievement	452 (443, 462)	515 (503, 528)	<0.01
Science achievement	462 (51, 473)	530 (518, 543)	<0.01
Adjusted model (with covariate controls)			
Instructional support			
Autonomy-support	−0.057 (−0.077, −0.036)	0.046 (0.030, 0.061)	<0.01
Instrumental help	−0.073 (−0.101, −0.036)	0.001 (−0.010, 0.011)	<0.01
Teacher warmth	−0.046 (−0.065, −0.028)	0.094 (0.070, 118)	<0.01
School-average achievement			
Reading achievement	498 (487, 508)	537 (529, 545)	<0.01
Mathematics achievement	478 (461, 495)	511 (496, 526)	<0.01
Science achievement	503 (484, 523)	539 (521, 557)	<0.01

The achievement values are reported as whole numbers as per OECD guidelines (OECD, 2019b). For comparison, OECD means and SD for the broader Australian sample that participated in PISA 2018 are, respectively, 503 and 105 for reading achievement, 491 and 85 for mathematics achievement, and 503 and 95 for science achievement. The instructional support factors were directly estimated in standardized units across the whole sample with $M \approx 0$, $SD = 0.094$ for autonomy-support, $SD = 0.089$ for instrumental help, and $SD = 0.135$ for teacher warmth.

not previously identified. Importantly, these six profiles were identical in both the Australian and English samples. In Phase 2 (the extension phase, conducted among a subsample of Australian teachers and schools), our results revealed that schools including a greater proportion of low-SES students were more likely to correspond to the Unsupportive school profile. In addition, the Supportive school profile was found to be associated with significantly higher school-average levels of autonomy-support, instrumental help, warmth, and achievement in reading, mathematics, and science. Key findings and implications are discussed below. Because many of the findings from Collie et al. (2020a) were replicated, we focus on the findings that are different, along with those that extend prior work.

Findings of Note From Single-Level Analyses in Phase 1

Six teacher profiles, identical across Australia and England, were identified in this study. Four of these profiles replicated Collie et al.'s (2020a) results (i.e., Low-Demand-Flourisher, Mixed-Demand-Flourisher, Job-Resourced-Average, and Balanced-Average profiles). Collie et al.'s (2020a) Struggler

profile was also identified in the present study, but had to be re-labeled the Low-Resource-Struggler profile to reflect the fact that the present results revealed one additional struggler profile, labeled the Mixed-Resourced-Struggler profile. This new profile was characterized by high job demands, mixed job resources (average teacher collaboration, high input), and low self-efficacy. In that sense, this new profile was similar to the Low-Resourced-Struggler profile in some regards (both characterized by high demands and low self-efficacy), but differed by displaying a rather high (vs. low) level of input in decision-making and a close to average (vs. very low) level of teacher collaboration.

It is possible that this new profile might have emerged as a result of the changes to the nature of teachers' work that have occurred between 2013 and 2018. For example, recent policy priorities have highlighted the need to better support teachers (OECD, 2019a), which might have enabled some of the more at-risk teachers to have greater input at work, and to benefit from slightly improved collaboration opportunities. Indeed, there have been efforts to seek teachers' perspectives on workload and to reduce unnecessary tasks (Higton et al., 2017; OECD, 2019a), which might have been particularly helpful to at least a subset of the teachers who struggled the most with their job demands. At the same time, the complexity of teachers' work has also received greater acknowledgment recently (Guerriero and Révai, 2017). Thus, the Mixed-Resourced-Struggler profile might both reflect the increased awareness of the need to support teachers and to involve them in decision making, but also the growing pressures of the job. Future research, including qualitative and mixed methods approaches, will be helpful to better understand the contingencies at play in the emergence of this new profile.

In relation to the relative prevalence of the identified profiles, it is interesting to consider differences across our study and prior results. Collie et al. (2020a) found that the two Flourisher profiles represented about one-third of their combined sample, the two Average profiles represented about one-half, and the Struggler profile represented one-fifth. In our study, the proportions were similar for the two Flourishers profiles (about one-quarter). However, the proportions were the opposite to those found by Collie et al. (2020a) for the two Average profiles (about one-quarter, rather than one-half) and for the two Struggler profiles (about one-half, rather than one fifth). Thus, although the profiles were similar in nature across the two studies, their prevalence differed greatly. Perhaps these shifts may be explained by some of the policy and practice changes that occurred between 2013 and 2018. For example, increases in compliance, accountability, and external evaluations across both countries between 2013 and 2018 (OECD, 2019a) may have meant teachers had less time and energy for professional development, managing disruptive behavior, and collaborating with colleagues. These changes, in turn, may have led to the greater prevalence of the two Struggler profiles in 2018. Compliance and accountability are also known to be negatively associated with reduced self-efficacy among teachers (von der Embse et al., 2016). Taken together, a key strength of the current study was the use of a replication sample collected 5 years after the original study. More precisely, it was because we conducted replication with nationally-representative data that we were able to identify teacher (and school) demand-resource profiles and then consider how macro-level changes to policy and practice may be implicated in the workplace experiences of distinct subpopulations of teachers. Going forward, it will be important to extend this knowledge by collecting data from the same teachers longitudinally to test whether teachers move between profiles over time, and what initiatives help teachers to move into more adaptive profiles.

The results involving the predictors and outcomes at the teacher-level were largely the same as those reported by Collie et al. (2020a), with the exception that, in addition to obtaining predictive similarity (i.e., evidence that predictions were the same across counties) like Collie et al. (2020a), we also obtained evidence of explanatory and adjusted-explanatory similarity (i.e., evidence that associations with outcomes where the same across countries before, and after controlling for school characteristics). Thus, the Mixed-Demand-Flourisher typically displayed the highest levels of job satisfaction and commitment, while the Low-Resourced-Struggler displayed the lowest levels. In practical terms, these results suggest that there might be merit in adopting a broad focus on reducing demands and increasing resources. This could involve the development of professional learning communities to help teachers build positive collaboration and self-efficacy (e.g., Durksen et al., 2017). Inviting teachers to have a say in school-level decisions and actively listening to teachers' perspectives and needs may be helpful for building their input in decision-making (see Collie et al., 2020a for additional implications for practice).

Findings of Note From Multilevel Analyses in Phase 1

Like Collie et al.'s (2020a), the present study led to the identification of a Supportive school profile and of an Unsupportive school profile that were equivalent across Australia and England. The results related to the outcomes of these profiles at the school level also generally match those reported, and discussed, by Collie et al. (2020a). For this reason, we focus on several novel findings related to the predictors of membership into these profiles at the school level. Notably, our results involving predictors provide the first source of information regarding associations between school characteristics and school demand-resource profiles. More precisely, our results showed that there was equivalence in the way these predictors were associated with the school profiles across the two countries. Thus, in Australia and England, schools with a higher proportion of low-SES students were more likely to correspond to the Unsupportive school profile than to the Supportive one. This finding is, unfortunately, unsurprising given that schools serving disadvantaged students often face many additional challenges and are often under resourced (Thomson et al., 2020), which is likely to result in higher demands and fewer resources for teachers. Going forward, it will be important to examine school funding and resourcing alongside SES

to disentangle their different roles. More broadly, funding cuts and austerity measures that have occurred over the past few years to services beyond school (e.g., services for disadvantaged families) have also likely increased the challenges for teachers and schools serving more disadvantaged students (e.g., Hanley et al., 2020). These broader societal impacts are important to consider in future research. Notably, school location, school size, and the proportion of students from non-English speaking backgrounds did not significantly predict the school profiles. Thus, beyond the role of SES, these other factors were not found to be associated with demand-resource profiles.

Findings of Note Involving Student Outcomes in Phase 2

Our final set of findings involved associations between the school profiles and the student outcomes, which has not been examined in prior research. More precisely, these results showed that, when compared to the Unsupportive school profile, the Supportive school profile tended to be associated with higher levels of school-average autonomy-support, instrumental help, and warmth, as well with higher levels of student achievement (across all three indicators of achievement considered in this study). These findings provide new evidence that demand-resource profiles matter not only for teachers' outcomes, but also for students' outcomes. The results extend prior work, which has demonstrated that teacher personal resource profiles tended to be associated with instructional support at the classroom-level (Klusmann et al., 2008), by showing that these associations are also salient at the school-level. Indeed, both the prevalence and nature of different teacher profiles within a school are associated with student outcomes, and this occurs even after controlling for school characteristics like SES. It is possible that these findings occurred because schools in which a greater number of teachers experience positive relationships at work, experience greater support at work, and are more confident in their teaching are likely to afford a more supportive and effective learning environment for students (e.g., Burić and Kim, 2020). More precisely, when teachers at a school feel supported and confident, they are more likely to create an environment where students feel better supported and are enriched in their learning (Collie et al., 2020a). In practical terms, these findings highlight the importance of creating supportive working climates not only for teachers, but also for students.

Limitations

Several limitations are important to consider when interpreting our study's findings. First, although the use of TALIS 2018 data comes with significant strengths (adequate sample size to conduct multilevel modeling, nationally representative data), it remains cross-sectional in nature. This means that we were not able to test for the directionality of any of the associations between the profiles and the outcomes. Importantly, our study was grounded in theory, which provided support for our hypotheses. Notwithstanding this, going forward it will be interesting for longitudinal and intervention research to examine the extent to which there may be reciprocal relations among the profiles and outcomes. Second, a strength of our study was that it involved data from both teachers and students. However, the TALIS-PISA link does not allow us to match students and teachers at the classroom-level. Thus, we cannot know whether the students who participated in PISA were taught by the teachers who participated in TALIS. Importantly though, all the students and teachers in our study were members of the same school—and that is where we focused in linking students and teachers (i.e., at the school-level). In future, it will be worth expanding this to consider links at the classroom-level, and also to include principal data about the job demands and resources present in the school. Given that the TALIS-PISA 2018 link was not conducted for England, it will also be important to test whether and how the links between the profiles and the student outcomes can be replicated in England, as well as in other countries. Third, one potential criticism of person-centered analysis is whether profiles are idiosyncratic to a particular sample. Our analytic approach has several important strengths that help to address this concern: profiles were examined (and found to be equivalent) across two countries and also largely replicated prior research (Collie et al., 2020a). Other strengths of our approach were that profiles were examined at two levels, and predictors and outcomes were entered after profile identification (and thus did not influence the profile solutions). Moreover, our models allowed us to examine important theoretical moderating mechanisms in more complex ways than possible in variable-centered modeling. Notwithstanding these strength, additional research (including more replication) is needed to provide additional evidence about the generalizability of our profiles. Because our study was intended to act as a replication and an extension, we focused on Australia and England to align with the Collie et al. (2020a) study. Given that the demands and resources examined in the present study are also experienced by teachers worldwide, in future it will be important to ascertain whether similar profiles can be identified in other countries as well.

CONCLUSION

The aim of the present study was to conduct a replication and extension of prior research examining demand-resource profiles among teachers and schools. Many of the findings from Collie et al. (2020a) were replicated, providing important support for the profiles and their relevance to teachers in Australia and England, and potentially in other contexts. In addition, several novel findings help to advance knowledge about the role of demands and resources in teachers' work and how these are linked not only with teachers' outcomes, but also students' outcomes. We speculate that some notable changes experienced within educational systems between these two studies might have had an impact on the current results, thus highlighting the importance of monitoring the impact of political and administrative changes for different educational systems worldwide. Moreover, our results make it clear that teachers should be supported at work. Not only

is this important for the individual teacher, it is relevant for shaping the working and learning climate across a whole school.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.oecd.org/education/talis/.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the UNSW Human Research Ethics Committee. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

RC would like to thank the Australian Research Council for funding this research (DE200100436). AJSM was supported by a grant from the Social Science and Humanity Research Council of Canada (435-2018-0368) in the preparation of this paper.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.711173/full#supplementary-material

REFERENCES

Ang, R. P. (2005). Development and validation of the teacher-student relationship inventory using exploratory and confirmatory factor analysis. *J. Exp. Educ.* 74, 55–74. doi: 10.3200/JEXE.74.1.55-74

Asparouhov, T. (2005). Sampling weights in latent variable modeling. *Struct. Equ. Model.* 12, 411–434. doi: 10.1207/s15328007sem1203_4

Bakker, A. B., and Demerouti, E. (2017). Job demands–resources theory: taking stock and looking forward. *J. Occup. Health Psychol.* 22, 273–285. doi: 10.1037/ocp0000056

Bakker, A. B., Hakanen, J. J., Demerouti, E., and Xanthopoulou, D. (2007). Job resources boost work engagement, particularly when job demands are high. *J. Educ. Psychol.* 99, 274–284. doi: 10.1037/0022-0663.99.2.274

Bliese, P., Maltarich, M., and Hendricks, J. (2018). Back to basics with mixed-effects models: nine take-away points. *J. Bus. Psychol.* 33, 1–23. doi: 10.1007/s10869-017-9491-z

Broadley, T. (2010). Digital revolution or digital divide: will rural teachers get a piece of the professional development pie? *Educ. Rural Aust.* 20:63. Available online at: https://www.spera.asn.au/research/journals/

Burić, I., and Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: an analysis using multilevel structural equation modeling. *Learn. Instr.* 66:101302. doi: 10.1016/j.learninstruc.2019.101302

Burns, E. C., Martin, A. J., and Collie, R. J. (2019). Examining the yields of growth feedback from science teachers and students' intrinsic valuing of science: implications for student- and school-level science achievement. *J. Res. Sci. Teach.* 56, 1060–1082. doi: 10.1002/tea.21546

Collie, R. J. (2021). A multilevel examination of teachers' occupational commitment: the roles of job resources and disruptive student behavior. *Soc. Psychol. Educ.* 24, 387–411. doi: 10.1007/s11218-021-09617-y

Collie, R. J., Granziera, H., and Martin, A. J. (2018). Teachers' perceived autonomy support and adaptability: an investigation employing the job demands-resources model as relevant to workplace exhaustion, disengagement, and commitment. *Teach. Teach. Educ.* 74, 125–136. doi: 10.1016/j.tate.2018.04.015

Collie, R. J., Granziera, H., Martin, A. J., Burns, E. C., and Holliman, A. J. (2020b). Adaptability among science teachers in schools: a multi-nation examination of its role in school outcomes. *Teach. Teach. Educ.* 90:103148. doi: 10.1016/j.tate.2020.103148

Collie, R. J., Malmberg, L.-E., Martin, A. J., Sammons, P., and Morin, A. J. S. (2020a). A multilevel person-centered examination of teachers' workplace demands and resources: links with work-related well-being. *Front. Psychol.* 11:626. doi: 10.3389/fpsyg.2020.00626

Collie, R. J., and Martin, A. J. (2017). Adaptive and maladaptive work-related motivation among teachers: a person-centered examination and links with well-being. *Teach. Teach. Educ.* 64, 199–210. doi: 10.1016/j.tate.2017.02.010

Dicke, T., Stebner, F., Linninger, C., Kunter, M., and Leutner, D. (2018). A longitudinal study of teachers' occupational well-being: applying the job demands-resources model. *J. Occup. Health Psychol.* 23, 262–277. doi: 10.1037/ocp0000070

Durksen, T. L., Klassen, R. M., and Daniels, L. M. (2017). Motivation and collaboration: the keys to a developmental framework for teachers' professional learning. *Teach. Teach. Educ.* 67, 53–66. doi: 10.1016/j.tate.2017.05.011

Enders, C. K. (2010). *Applied Missing Data Analysis*. New York, NY: Guilford.

Fackler, S., Sammons, P., and Malmberg, L.-E. (2020). A comparative analysis of predictors of teacher self-efficacy in student engagement, instruction and classroom management in Nordic, Anglo-Saxon and East and South-East Asian countries. *Rev. Educ.* 9, 203–239. doi: 10.1002/rev3.3242

Guerriero, S., and Révai, N. (2017). "Knowledge-based teaching and the evolution of a profession," in *Pedagogical Knowledge and the Changing Nature of the Teaching Profession,* ed S. Guerriero (Paris: OECD), 253–269.

Hamre, B. K., Pianta, R. C., Downer, J. T., Decoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: testing a developmental framework of effective teaching in over 4,000 classrooms. *Elem. Sch. J.* 113, 461–487. doi: 10.1086/669616

Hanley, T., Winter, L. A., and Burrell, K. (2020). Supporting emotional well-being in schools in the context of austerity: an ecologically informed humanistic perspective. *Br. J. Educ. Psychol.* 90, 1–18. doi: 10.1111/bjep.12275

Higton, J., Leonardi, S., Richards, N., Choudhoury, A., Sofroniou, N., and Owen, D. (2017). *Teacher Workload Survey 2016*. UK Department for Education. Available online at: https://www.gov.uk/government/publications/teacher-workload-survey-2016

Keller, M. M., Hoy, A. W., Goetz, T., and Frenzel, A. C. (2016). Teacher enthusiasm: reviewing and redefining a complex construct. *Educ. Psychol. Rev.* 28, 743–769. doi: 10.1007/s10648-015-9354-y

Klassen, R. M., and Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychol.* 102, 741–756. doi: 10.1037/a0019237

Klassen, R. M., Perry, N. E., and Frenzel, A. C. (2012). Teachers' relatedness with students: an underemphasized component of teachers' basic psychological needs. *J. Educ. Psychol.* 104, 150–165. doi: 10.1037/a0026253

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., and Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: the important role of self-regulatory patterns. *J. Educ. Psychol.* 100, 702–715. doi: 10.1037/0022-0663.100.3.702

Litalien, D., Gillet, N., Gagné, M., Ratelle, C. F., and Morin, A. J. S. (2019). Self-determined motivation profiles among undergraduate students: a robust test of profile similarity as a function of gender and age. *Learn. Individ. Differ.* 70, 39–52. doi: 10.1016/j.lindif.2019.01.005

Mäkikangas, A., Tolvanen, A., Aunola, K., Feldt, T., Mauno, S., and Kinnunen, U. (2018). Multilevel latent profile analysis with covariates: identifying job characteristics profiles in hierarchical data as an example. *Organ. Res. Methods* 21, 931–954. doi: 10.1177/1094428118760690

McLarnon, M. J. W., and O'Neill, T. A. (2018). Extensions of auxiliary variable approaches for the investigation of mediation, moderation, and conditional effects in mixture models. *Organ. Res. Methods* 21, 955–982. doi: 10.1177/1094428118770731

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge.

Morin, A. J. S., and Litalien, D. (2017). *Longitudinal Tests of Profile Similarity and Latent Transition Analyses*. Montreal, QC: Substantive Methodological Synergy Research Laboratory. Retrieved from: http://smslabstats.weebly.com/uploads/1/0/0/6/100647486/lta_distributional_similarity_v02.pdf

Morin, A. J. S., and Litalien, D. (2019). "Mixture modeling for lifespan developmental research," in *Oxford Research Encyclopedia of Psychology*. New York, NY: Oxford University Press. doi: 10.1093/acrefore/9780190236557.013.364

Morin, A. J. S., Meyer, J. P., Creusier, J., and Biétry, F. (2016). Multiple-group analysis of similarity in latent profile solutions. *Organ. Res. Methods* 19, 231–254. doi: 10.1177/1094428115621148

Muthén and Muthén (2019). *Mplus User's Guide*. Los Angeles, CA: Muthén and Muthén.

OECD (2009). *Creating Effective Teaching and Learning Environments: First Results From TALIS*. OECD. Retrieved from: http://www.oecd.org/education/school/ doi: 10.1787/9789264068780-en

OECD (2014). *TALIS* 2013 *Technical Report*. Paris: OECD. Retrieved from: http://www.oecd.org/education/school/

OECD (2019a). *Education Policy Outlook 2019: Working Together to Help Students Achieve Their Potential*. OECD. doi: 10.1787/2b8ad56e-en

OECD (2019b), *PISA 2018 Assessment and Analytical Framework*. OECD.

OECD (2019c). *TALIS* 2018 *and TALIS Starting Strong* 2018 *User Guide*. OECD. Retrieved from: http://www.oecd.org/education/talis/

OECD (2020). *PISA* 2018 *Technical Report*. OECD. Retrieved from: https://www.oecd.org/pisa/data/pisa2018technicalreport/

Parliament of Australia (2019). *Inquiry Into the Status of Teaching: Summary*. Retrieved from: https://www.aph.gov.au/

Raykov, T., and Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parameter functions in SEM. *Struct. Equat. Model. Multidiscip. J.* 11, 621–637. doi: 10.1207/s15328007sem1104_7

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York, NY: Wiley.

Ryan, R. M., and Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. New York, NY: Guilford Press.

Simbula, S., Panari, C., Guglielmi, D., and Fraccaroli, F. (2012). Teachers' well-being and effectiveness: the role of the interplay between job demands and job resources. *Proc. Soc. Behav. Sci.* 69, 729–738. doi: 10.1016/j.sbspro.2012.11.467

Sizmur, J., Ager, R., Bradshaw, J., Classick, R., Galvis, M., Packer, J., et al. (2019). *Achievement of 15-Year-Olds in England: PISA* 2018 *Results*. UK Department for Education. Available online at: http://www.gov.uk/government/publications

Skaalvik, E., and Skaalvik, S. (2018). Job demands and job resources as predictors of teacher motivation and well-being. *Soc. Psychol. Educ.* 21, 1251–1275. doi: 10.1007/s11218-018-9464-8

Skinner, E. A., and Belmont, M. J. (1993). Motivation in the classroom: reciprocal effects of teacher behavior and student engagement across the school year. *J. Educ. Psychol.* 85, 571–581. doi: 10.1037/0022-0663.85.4.571

Thomson, S., De Bortoli, L., Underwood, C., and Schmid, M. (2020). *PISA 2018 Reporting Australia's Results: Volume II Student and School Characteristics*. Australian Council for Educational Research. Available online at: https://research.acer.edu.au/ozpisa/49/

Tschannen-Moran, M., and Woolfolk Hoy, A. E. (2001). Teacher efficacy: capturing an elusive construct. *Teach. Teach. Educ.* 17, 783–805. doi: 10.1016/S.0742-051X.(01)00036-1

UK Department for Education (2018). *Making Data Work: Report of the Teacher Workload Advisory Group*. Retrieved from: https://www.gov.uk/government/groups/teacher-workload-advisory-group

Van den Broeck, A., De Cuyper, N., Luyckx, K., and De Witte, H. (2012). Employees' job demands–resources profiles, burnout and work engagement: a person-centred examination. *Econ. Indust. Democr.* 33, 691–706. doi: 10.1177/0143831X11428228

Vermunt, J. K. (2010). Latent class modeling with covariates: two improved three-step approaches. *Polit. Anal.* 18, 450–469. doi: 10.1093/pan/mpq025

von der Embse, N. P., Pendergast, L. L., Segool, N., Saeki, E., and Ryan, S. (2016). The influence of test-based accountability policies on school climate and teacher stress across four states. *Teach. Teach. Educ.* 59, 492–502. doi: 10.1016/j.tate.2016.07.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 24 September 2021
doi: 10.3389/feduc.2021.744347

The Influence of School Context Factors on the Induction Support as Perceived by Newly Qualified Teachers

*Marieke Van der Pers and Michelle Helms-Lorenz**

Department of Teacher Education, University of Groningen, Groningen, Netherlands

This study explores the relationship between school demographic characteristics and the amount of provided induction activities, as perceived by beginning teachers (BTs), with special attention for Professional Development Schools and non-Professional Development Schools. The aim is to provide information that is useful to improve induction arrangements to particular school contexts. Data were collected in a Dutch national induction program in which qualified BTs are supported in their first 3 years of professional practice. The support monitor measures multiple induction activities with regard to the implementation of workload reduction, school enculturation, professional development plans and lessons support. 1,670 BTs working in 195 schools reported on these activities. Linear regression analyses revealed that less induction activities were associated with schools with multiple locations, more enrolled students and number of BTs employed, and with schools with older male teaching staff. No significant differences were found between the amount of support perceived by BTs in PDSs and non-PDSs. Yet school characteristics revealed stronger predictive values for the amount of support provided in non-PDS settings. These insights are relevant for mapping school differences in induction arrangements in order to enhance the equity of support across schools to assure the development of teaching skills of BTs.

Keywords: induction, support, beginning teachers, school context, Netherlands

OPEN ACCESS

Edited by:
Antonio P. Gutierrez de Blume, Georgia Southern University, United States

Reviewed by:
Balwant Singh, Partap College of Education, India
Luxin Yang, Beijing Foreign Studies University, China

****Correspondence:***
Michelle Helms-Lorenz
m.helms-lorenz@rug.nl

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *20 July 2021*
Accepted: *06 September 2021*
Published: *24 September 2021*

Citation:
Van der Pers M and Helms-Lorenz M (2021) The Influence of School Context Factors on the Induction Support as Perceived by Newly Qualified Teachers. Front. Educ. 6:744347. doi: 10.3389/feduc.2021.744347

INTRODUCTION

Within the teaching community teacher induction is a widespread professional development intervention aimed to attract, develop and commit beginning teachers (BTs) to teaching. The work of Ingersoll and Smith (2004) provides valuable guidelines that lead to higher teacher performance and higher retention rates. Teacher induction arrangements based on these guidelines have been implemented in the Netherlands. An experimental study applying a 3 year induction arrangement showed a longitudinal positive effect on teaching quality of BTs, as perceived by students (Maulana et al., 2015; Helms-Lorenz et al., 2016). In addition, this programme had a positive effect on teacher retention (Helms-Lorenz et al., 2016).

In the period 2013–2019 the Dutch Ministry of Education supported the implementation of a national project aimed to implement induction arrangements in at least 50% of secondary schools. This was done according to a pre-defined framework in which a set of coherent support activities in four key-induction elements were combined and carefully timed to the needs of BTs, guided by

regional teacher education institutes. The four elements concern 1) providing appropriate workload; 2) introducing the school culture and school policy; 3) implementation of professional development plans; and 4) providing classroom support in the form of feedback and coaching based on classroom observations and preparing and discussing lessons with a coach (Helms-Lorenz et al., 2015; Helms-Lorenz et al., 2018a). The 447 schools participating in this project were closely supported by teacher education institutes to help tailor the implementation of the framework to each school's specific social, political, cultural, individual, organizational and administrative context. This collaborative approach, however, does not guarantee that all participating schools were able to introduce and retain induction arrangements for BTs effectively. This study aims to get insight into the extent to which this large scale intervention has succeeded in schools offering BTs an adequate amount of support activities, regardless their contextual demographic characteristics. These demographic characteristics are school size, age and gender composition of the teaching staff, denomination, and the socioeconomic composition of the student population. As existing support- and learning infrastructures are assumed to be more organized in schools that have a formal collaboration relationship and structure with a teacher education institute, a distinction is made between Professional Development Schools (PDSs) and non-Professional Development Schools (non-PDSs). The findings may reveal that the national induction framework should be tailored more specifically to certain school context factors in order to reach all targeted BTs.

Significance of Induction Arrangements for Beginning Teachers

Induction arrangements aim to enhance teaching quality and to commit BTs to the teaching profession. The work of Ingersoll and Smith (2004) has had a great impact on the conceptualization and development of induction arrangements across the world. Their contribution to the field is especially important because they unravelled relative contributions to teacher retention of different features of induction arrangements. In their large scale survey study (N = 3,235), the strongest activities contributing to retention were having a mentor from the same field, having common planning time with other teachers in the same subject, having regular scheduled collaboration with other teachers, and being part of an external network of teachers. They also found that, as the number of components in the induction arrangement increased, the probability of teacher turnover decreased.

In addition to the association with teacher retention, a number of studies have shown positive effects of induction arrangements and high quality mentoring programs on job satisfaction and commitment (Richardson et al., 2007; Ingersoll and Strong, 2011), improved classroom instruction and student achievement (Ingersoll and Strong, 2011). In the Dutch context, a multicomponent induction arrangement on perceived teaching quality reported by students, revealed positive improvements of over time (Maulana et al., 2015; Helms-Lorenz et al., 2016). Researchers identified the separate support elements responsible for this effect and found that a low degree of workload reduction and enculturation in the school contributed to teacher attrition and that classroom support had a positive impact on the development of teaching skills. The same research group also showed that specific induction arrangement elements appear to be meaningful to reduce the level of specific perceived stress causes and stress responses. Harmsen et al. (2018) found that workload reduction decreased the level of perceived high psychological task demands, negative social aspects, and stress responses. Classroom support decreased the level of perceived negative emotions and discontent, and school enculturation decreased perceived discontent over time. Similarly, in the Dutch context, Kessels (2010) found a positive effect of induction arrangements on BTs' well-being in general.

School Context and the Effective Implementation of Induction Arrangements

According to Firestone and Louis (1999), the formal structures that belong to an induction arrangement are effective when they are in line with, and reinforced by, the norms, values, and practices of the professional culture in which they are embedded. An effective professional culture is characterized by malleable conditions such as school leadership, school policies, and organizational conditions (Creemers and Kyriakides, 2010; Kraft and Papay, 2014; Kutsyuraba et al., 2016; Scheerens, 2016). Birkeland and Feiman-Nemser (2012) note that the success of a school-based induction arrangement relies on the commitment and investment of school leaders, and on the entire school taking responsibility to help new teachers succeed. Examples of factors that contribute to the variation in induction implementation and teacher experiences are leadership characteristics, clarity of tasks and responsibilities for beginning teacher support, types and formats of support, formal and informal rules, and organizational framework (Kutsyuruba et al., 2016). Ideally, induction arrangements are adapted to such school organizational conditions in the design and implementation phase. As these factors are complex, multi-layered, and hard to measure objectively, it is difficult to fully capture them within a research design. As these objective characteristics are associated indirectly with the organisational contexts of schools (Scheerens, 2016), this study aims to provide an indication of which school demographic characteristics facilitate or hinder the implementation of induction activities.

School Size and Proportion of Beginning Teachers

School size is an important factor influencing the number and type of programs schools offer to teachers and it also impacts the social interactions among the staff. In small schools, the social network is more compact, which enhances (informal) interpersonal relationships, teacher collaboration, and close relationships between teachers and students. In such contexts, staff are more likely to engage in common activities and teachers have greater autonomy (Fowler and Walberg, 1991; Lee et al.,

1993; Maslowski, 2001; Opdenakker and Van Damme, 2007). In larger school organisations, it is likely that more formal rules and procedures set the basis for activities; relationships may be more formal and staff members may have less autonomy (Lee et al., 1993; Maslowski, 2001). This implies that school size could influence the way support is provided to and experienced by BTs.

Efficiency in designing and implementing induction arrangements may be different in schools with a larger proportion of BTs in their staff than in schools that have relatively few BTs. For one, providing support activities to a larger group of BTs requires a larger number of supportive staff and demands a more detailed supportive structure for professionalization activities. We explore whether BTs working in schools with different student numbers and with different proportions of BTs employed, perceive different amounts of induction activities. The direction of the association between school size, proportion of BTs and induction effects is unclear, as both small and large scale situations can contain factors facilitating or hindering collaboration and supportive structures.

Age and Gender Composition of Teaching Staff

Although no studies were found focusing on the association between the age composition of schools' teaching staff and the provided support offered to BTs, researchers have studied the effect of age composition of teaching staff on learning cultures. According to Day and Gu (2007), schools may encounter difficulties in meeting support needs all staff members as they vary in professional, personal, and workplace conditions during different stages of the teaching career, which this affects teachers' commitment and resilience in different ways. For instance, BTs may prioritise support by head teachers and colleagues and a supportive school culture for their professional and identity development, while midcareer teachers face tension in managing professional and personal lives and are in need of support from trustable head teachers and colleagues, and veteran teachers are challenged by adjusting to (educational) change and health problems (Day and Gu, 2007). Kardos et al. (2001) studied the association between age composition of teaching staff and the type of learning cultures for BTs. They state that "a new teacher's encounter with professional culture will depend on the group of colleagues with whom she works, how they interact, and whether they welcome novices in their professional exchanges and pay attention to their needs and concerns". They describe different learning cultures in schools where the age composition of the teaching staff is predominantly young (novice-oriented professional cultures) or older (veteran-oriented cultures). Due to the lack of research on school employee composition, there is little insight in how this factor is associated with the type and amount of induction activities for BTs.

In addition to age, the gender composition of a school's staff can affect the social dynamics within schools. In general, women are more interested in informal communication and are more hesitant towards individualistic work settings (Bryk et al., 1999). As this is a rather complex and debated issue in the organizational literature, in this study we account for the gender composition of a teaching staff in order to explore the possible association with the provision of induction activities to BTs.

Denomination of Schools

The cultural identity of a school is another characteristic that may influence the type and number of support activities organized for (beginning) teachers. Within a school system, historical norms and values differ by school denomination. Maslowski (2001) investigated the association between school culture and the performance of Dutch secondary schools and showed that Dutch secondary schools are rather similar with regard to values like collaboration, commitment and stability, but found substantive differences when comparing these orientations among denominations. Schools without a religious or philosophic denomination tend to value innovation, insight and adaptation and are therefore more inclined towards educational change than Catholic or Protestant schools. Protestant schools tend to value stability and continuity and are therefore more reluctant to change and to implement innovations and (Maslowski, 2001). Specific elements of the induction arrangement may be valued and implemented in a differently by a school's denomination, expecting to lead to differences in the amount of perceived support by BTs.

Socioeconomic Composition of Students

International research has shown that the socioeconomic background of students is associated with the policies, practices and resources of schools, and thereby the context in which (beginning) teachers are working (Lee et al., 1993; Loeb et al., 2005). Specific teaching skills are required to deal with the relatively high proportion of students with learning difficulties and behavioral problems and low socioeconomic backgrounds (Muijs et al., 2004; Sykes and Musterd, 2011; Sykes and Kuyper, 2013). In such challenging contexts, (beginning) teachers are less qualified (Bonesrønning et al., 2005; Boyd et al., 2005; Danhier, 2016; Lankford et al., 2002; Organisation for Economic Co-operation and Development Organisation for Economic Co-operation and Development, 2018, less experienced (Sass et al., 2012; Danhier, 2016), learn less (Lankford et al., 2002; Hanushek et al., 2004; Ronfeldt et al., 2016), have lower retention rates (Scafidi et al., 2007; Danhier, 2016), and cooperate less (Opdenakker and Van Damme, 2001; Sass et al., 2012). Schools with a more challenging student population are therefore in greater need for providing adequate teacher support. For the US context, Johnson et al. (2004) found a support gap between high- and low-income schools, where novice teachers working in low-income schools were supported less. The Dutch educational system has fairly homogeneous school conditions regarding teacher salaries and teacher qualification criteria, however some schools do have a high proportion of students with a lower socioeconomic backgrounds (Organisation for Economic Co-operation and Development, 2016). It is not clear to what extent the above mentioned features are associated with perceived support. This study accounts for the composition of the SES of a school's student population in order to get insight into the possible association with the provision of induction activities to BTs.

TABLE 1 | School characteristics and reported support activities, national and study sample, stratified by PDSs and non-PDSs.

	National sample (N = 651)			Study sample (N = 195)			PDSs (N = 89)			Non-PDSs (N = 106)		
	%	*Mean*	*SD*	*%*	*Mean*	*SD*	*%*	*Mean*	*SD*	*%*	*Mean*	*SD*
Schoolsize	—	—	—	—	—	—	—	—	—	—	—	—
One location	57.0	—	—	50.8	—	—	46.1	—	—	54.7	—	—
Multiple locations <2,500 students	25.4	—	—	23.1	—	—	20.2	—	—	25.5	—	—
Multiple locations>=2,500 students	17.7	—	—	26.2	—	—	33.7	—	—	19.8	—	—
Proportion beginning teachers (proxy)	—	—	—	—	—	—	—	—	—	—	—	—
≤5%	n.a	—	—	42.1	—	—	42.7	—	—	41.5	—	—
5–10%	n.a	—	—	34.4	—	—	32.6	—	—	35.8	—	—
>10%	n.a	—	—	23.6	—	—	24.7	—	—	22.6	—	—
Age and gender composition teaching staff	—	—	—	—	—	—	—	—	—	—	—	—
Average age male teachers		46.8	3.1		46.5	2.6	—	—	2.3		46.7	2.9
Average age female teachers		42.8	3.0		42.4	2.6	—	—	2.2		42.4	2.8
Proportion of female teachers		52.9	7.5		52.1	5.5	—	—	5.3		51.9	5.7
Professional Development School	—	—	—	—	—	—	—	—	—	—	—	—
Yes	28.6	—	—	45.6	—	—	—	—	—	—	—	—
No	71.4	—	—	54.4	—	—	—	—	—	—	—	—
School identity		—	—		—	—	—	—	—	—	—	—
No religious/philosophic background	28.7	—	—	34.9	—	—	32.6	—	—	36.8	—	—
Catholic	23.8	—	—	27.2	—	—	36.0	—	—	19.8	—	—
Protestant	30.0	—	—	24.6	—	—	23.6	—	—	25.5	—	—
Philosophic background	17.5	—	—	13.3	—	—	7.9	—	—	17.9	—	—
Proportion SES neighbourhood students	—	—	—	—	—	—	—	—	—	—	—	—
Highest SES	—	25.8	21.9	—	31.0	23.0	—	29.6	20.7	—	32.1	24.9
Lowest SES	—	30.6	24.5	—	25.9	21.6	—	24.8	19.8	—	26.8	23.0
Degree of urbanisation municipality	—	—	—	—	—	—	—	—	—	—	—	—
Very urban	23.8	—	—	25.6	—	—	22.5	—	—	28.3	—	—
Urban	33.6	—	—	32.3	—	—	31.5	—	—	33.0	—	—
Suburban	18.6	—	—	23.1	—	—	27.0	—	—	19.8	—	—
Rural	22.7	—	—	19.0	—	—	19.1	—	—	18.9	—	—
Reported support activities	—	—	—	—	9.9	2.4	—	9.9	2.2	—	10.0	2.6
Workload reduction (max 19)	—	—	—	—	4.1	1.0	—	3.9	1.0	—	4.2	1.0
School enculturation (max 6)	—	—	—	—	6.1	1.4	—	6.0	1.3	—	6.1	1.5
Professional development (max 10)	—	—	—	—	5.7	1.5	—	5.7	1.5	—	5.7	1.5
Effective teaching behavior (max 9)	—	—	—	—	9.9	2.4	—	9.9	2.2	—	10.0	2.6

Existing Collaboration With Teacher Education Institutes

Currently, an increasing number of schools and teacher education institutes form partnerships to collaborate in the design and implementation of teacher education programs and to provide a strong support system in which they both supervise and support the development of pre-service teachers. Schools that participate in such partnerships are called Professional Development Schools (PDSs). In this approach, field experience is spread throughout the teacher preparation program, aiming to closely connect theory with practice.

A recent small-scale longitudinal study in the Netherlands revealed that student teachers who were educated in PDSs had higher teaching skills and reported more opportunities to learn than a similar group of teachers working at non-PDSs (Helms-Lorenz et al., 2018b). It is expected that induction arrangements for BTs benefit from these existing learning infrastructures, providing more support activities to BTs working in PDSs schools than those working in non-PDSs schools. As in PDSs the infrastructure for learning is more integrated in the school organization, it is expected that the demographic school demographic characteristics have a weaker predictive value in explaining differences between schools in providing support activities to BTs than they have in non-PDSs. As in this study the participating schools could be defined by being a PDS or not, it will be possible to investigate these expectations.

The Dutch School Context and the National Three-Year Induction Arrangement

In 2015 there were 651 registered secondary schools, with an average size of about 1,500 students per school (**Table 1**). The school education system is subsidized by the government and parents are free to choose any school they prefer. Secondary schools can be categorized as public schools (29%) and schools that are based on a particular religious (54%) or educational philosophy (17%) (Dienst Uitvoering Onderwijs, 2016). It is important to note that in the Netherlands public schools do not have a religious profile, and that schools with a religious profile are not very different from public schools.

More than half of all Dutch schools are located in more urbanized areas and nearly one third is located in the

economic core. Due to declining birth rates, the student population for secondary education is predicted to decrease with more than five percent in the period 2015–2020 (Dienst Uitvoering Onderwijs, 2015b; Van den Berg et al., 2015). In 2014, 30 percent of all secondary school students in the Netherlands were from neighbourhouds with the lowest socioeconomic status, and a quarter of all students were from the highest SES neighbourhoods (Dienst Uitvoering Onderwijs, 2015a; Sociaalen Cultureel Planbureau, 2014. According to the latest OECD figures, the Dutch school system is one of the best in the OECD in the sense that a relatively high proportion of students with a disadvantaged background succeeding at school (Organisation for Economic Co-operation and Development, 2016). There are however remarkably large differences between schools concerning student achievement that may be explained by learning climate, teacher quality and learning materials (Inspectie van het Onderwijs, 2017).

Nearly one third of the teacher population in the Netherlands is 55 years and older (Centraal Bureau voor de Statistiek, 2018). Although the gender composition of the full teacher population is almost equal (46% male and 54% female teachers), a larger proportion of the older teachers is male (58%), whereas the younger teachers are more often female (60%) (Centraal Bureau voor de Statistiek, 2018). Female teachers in particular prefer to work part time.

Teacher education programs lead to two types of degrees. A first level degree is obtained after graduation at the university (master) level, a second level degree is obtained at the bachelor level. Teachers with a first level degree are qualified to teach in the lower and upper tracks in secondary education. Teachers with a second level degree are qualified to teach in the lower tracks. Qualified BTs hold different positions at the labour market; in the first years the majority will have a temporary position, some a permanent position; being employed at one more schools; working full- or part-time. This work status is determined by a combination of possibilities and restrictions of the (regional) labour market and depends on individual preferences.

Since 2009, PDSs aim to bridge the gap between the job requirements and the theoretical curriculum requirements (Nederlands-Vlaamse Accreditatieorganisatie Nederlands-Vlaamse Accreditatieorganisatie, 2009. In these partnerships schools and teacher education institutes collaborate to educate student teachers and share the responsibility for the content and implementation of the curriculum. This implies an established professional learning infrastructure within the school.

Most schools in the Netherlands offer a 1 year induction arrangement for BTs. In this study, a 3 year induction program was implemented in 50% of all regions of the Netherlands, with the assistance of eight regional teacher education institutes to develop, implement, and evaluate the support activities. The implementation aimed to tailor the programme elements to the specific BT needs of each of the 447 participating schools. The framework combines a set of coherent support activities in four key-induction activities which are clearly described and carefully timed to the needs of BTs: 1) providing appropriate workload; 2) introducing the school culture and school policy; 3) implementation of professional development plans; and 4) providing classroom support in the form of feedback and coaching based on classroom observations and preparing and discussing lessons with a coach. A detailed description of the project can be found in Helms-Lorenz et al. (2015) and Helms-Lorenz et al. (2018a).

This study explores the relationship between school demographic characteristics and the amount of provided induction activities, as perceived by beginning teachers (BTs) and thereby aims to provide information that is useful to improve induction arrangements to particular school contexts.

Research Questions

1. What amount of support is provided by participating schools, as perceived by beginning teachers participating in this national induction arrangement?
2. To what extent do demographic characteristics of schools explain potential differences in the offered amount of support activities?
3. Are demographic characteristics of schools less associated with the amount of support activities offered by Professional Development Schools compared to non-Professional Development Schools?

METHOD

Participants, Sample and Procedure

The data were collected in three waves (cohorts) of the national longitudinal study in secondary education in the Netherlands (Helms-Lorenz et al., 2015; Helms-Lorenz et al., 2018a). During the academic years 2014 through 2019, 2,317 qualified BTs enrolled in the study. At the end of each academic year, a research questionnaire, the so called "support monitor" was used to measure BTs' perceptions of the induction activities provided by their schools. This monitor has been used previously (Helms-Lorenz et al., 2012; Harmsen et al., 2018a), details will be discussed in the next section. 2,040 teachers working at 286 schools filled in this questionnaire at least once. 149 cases were excluded as the teachers did not give consent to use the data for research purposes or due to incompleteness. For schools with multiple locations, at least five completed monitors were required to represent the offered induction activities at the school level. In order to capture the support activities provided to BTs working in smaller schools as well, all schools having one location were included in the study sample. This selection procedure resulted in a final sample of 195 schools among which1,670 BTs completed the support monitor at least once.

This sample consisted of nearly 30% of the total number of secondary schools in the Netherlands (**Table 1**) and characterizes an overrepresentation of schools with multiple locations and a larger student population, and of schools located in suburban areas. The average age and gender composition of the teaching staff did not deviate from national averages. Compared with the national sample, the study sample contains a larger proportion of PDSs, more public and Catholic schools, and more schools with a greater proportion of students living in higher SES neighborhoods.

Comparing the samples of PDS and non-PDS, a similar proportion of BTs, average age, and gender composition of the teaching staff (**Table 1**) is evident. Compared with non-PDSs,

PDSs more often have multiple locations, more students, are more often located in highly urbanized areas and less often in rural areas, slightly more often have a religious profile, and have a lower proportion of students living in the highest and lowest socioeconomic status neighborhoods.

Measures

Dependent Variables: Support Activities Provided to Beginning Teachers

In order to gain insight in the type and amount of support activities that a school offers BTs, all BTs completed the "support monitor". This monitor focuses on four specific elements of the national induction arrangement; facilities concerning workload reduction (18 items), activities concerning school enculturation (6 items), stimulating professional development (9 items), and supporting effective teaching behavior (9 items). Examples of questioned activities are: My schools offered me the opportunity to … "be exempted from teaching difficult classes", "meet with supervisors at set times", "work with a personal development plan", "to reflect with colleagues on professional development", "to evaluate classroom observations", and to "plan/prepare/execute my lessons with colleagues or my coach" (see **Supplementary Appendix S1A** for the questioned items for each of the four induction elements). All items were rated dichotomously (yes, was provided/no, was not provided) and for each individual teacher the nominal data was aggregated (counted) for each induction element. These aggregated teacher scores were used to calculate the mean perceived number of provided activities, per induction element, for each school.

Independent Variables: Demographic Characteristics of Schools

In the Netherlands, each registered school is represented by a unique administrative number (BRIN) that enabled us to retrieve contextual information which we connected to the data (Sociaalen Cultureel Planbureau, 2014; Dienst Uitvoering Onderwijs, 2015a; Dienst Uitvoering Onderwijs, 2016; Dienst Uitvoering Onderwijs, 2018).

The variable *school size* represents whether a school consists of a single or multiple locations and of the number of students being registered on October 1st, 2015. The three categories are 1) single location, 2) multiple locations with less than 2,500 students, 3) multiple locations with at least 2,500 students. Only three of the 99 single-location schools have at least 2,500 students.

Unfortunately, information about the exact number of BTs appointed per school was not available. The variable *proportion of beginning* teachers working at a school is therefore based on the number of BTs that the schools registered to participate in the national induction arrangement. As participating schools received financial support for each BT[1] enrolled in the project, creating an incentive to enrol all BTs, we assumed that the majority of BTs are captured with this proxy.

[1]In this project a beginning teacher is defined as a person certified to teach in secondary education, having less than 2 years of teaching experience after certification

The continuous variables *average age male/female teaching staff* represents the average age of a schools' teaching staff, as administered on October 1st, 2015 (Dienst Uitvoering Onderwijs, 2018). The continuous variable *proportion of female teachers* refers to the proportion of female teachers employed at a school, administered at the same date (Dienst Uitvoering Onderwijs, 2018).

The categorical variable *school identity* distinguishes the main four school types in the Netherlands by 1) public school, 2) Catholic denomination, 3) Protestant denomination, and 4) specific educational philosophy. These categories are in line with the work of Maslowski (2001).

Socioeconomic status of students' residential neighborhoods was used as a proxy for socioeconomic composition of schools' student populations. Information of the residential zip codes of schools' registered students (Dienst Uitvoering Onderwijs, 2015a) were merged to status scores of these neighborhoods (Sociaalen Cultureel Planbureau, 2014). These status scores are composed of average income, share of residents with a low income, share of residents with a low educational level, and share of unemployed residents (Sociaalen Cultureel Planbureau, 2014). The two continuous variables *proportion of highest/lowest SES students* represent the proportion of students living in neighborhoods with the highest and lowest socioeconomic status.

The variable *professional development school* indicates whether teachers were working at schools qualified as a Professional Development School by the Nederlands-Vlaamse Accreditatieorganisatie (NVAO) in the period 2009–2015 (see Nederlands-Vlaamse Accreditatieorganisatie, 2009 for qualification criteria).

Analytic Strategy

As none of the variables have missing values, no missing value analysis was required. Descriptive analyses provide insight in the amount of provided activities at school level (RQ1). Bivariate linear regression models were conducted to provide insight in the extent to which the described sociodemographic school characteristics predicted the amount of perceived activities (RQ2). General linear regression models including all independent variables were run separately for each of the four induction elements (RQ2). A multivariate model was not used, as the four elements can be implemented independently from each other. Interactions between the four dependent variables are therefore not useful for the interpretation of the results. The full models were stratified by PDSs and non-PDSs in order to gain insight into the extent to which the independent variables predict the offered amount of induction activities in each of these contexts (RQ3). All data preparation and analyses were conducted with SPSS Version 24.

RESULTS

Type and Amount of Perceived Support Activities

On average, the 195 schools offered about half of the measured activities concerning workload reduction and enculturation (**Table 1**). On average, schools offered about 60 percent of the activities concerning professional development and effective

TABLE 2 | Binary models: relation reported support activities by BTs and school characteristics, four induction elements, full sample (N = 195).

	Workload reduction			Enculturation			Professionalization			Effective teaching		
	Estimate	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*	*Estimate*	*SE*	*p*
Schoolsize[a]	—	—	—	—	—	—	—	—	—	—	—	—
Multiple locations <2,500 students	−0.858*	0.423	0.043	−0.254	0.178	0.152	−0.299	0.257	0.244	−0.118	0.270	0.663
Multiple locations>=2,500 students	−1.330**	0.406	0.001	−0.595***	0.170	0.000	−0.444	0.246	0.071	−0.113	0.258	0.662
Proportion of beginning teachers (proxy)[b]	—	—	—	—	—	—	—	—	—	—	—	—
5–10%	−1.318**	0.428	0.002	−0.147	0.186	0.431	−0.644*	0.257	0.012	−0.523	0.273	0.055
>10%	−1.829***	0.445	0.000	−0.333	0.194	0.085	−0.929***	0.267	0.001	−0.567*	0.284	0.046
Age and gender composition teaching staff	—	—	—	—	—	—	—	—	—	—	—	—
Average age male teachers	−0.042	0.066	0.519	−0.079**	0.027	0.003	−0.104**	0.038	0.007	−0.052	0.041	0.196
Average age female teachers	0.034	0.068	0.622	0.013	0.029	0.646	0.005	0.040	0.899	0.035	0.042	0.406
Proportion of female teachers	0.023	0.032	0.467	0.030*	0.013	0.025	0.025	0.019	0.175	0.030	0.020	0.124
Professional Development School[c]	—	—	—	—	—	—	—	—	—	—	—	—
Yes	−0.089	0.347	0.799	−0.283	0.145	0.051	−0.013	0.207	0.952	0.021	0.216	0.923
School identity[d]	—	—	—	—	—	—	—	—	—	—	—	—
Catholic	-0.586	0.441	0.184	-0.150	0.186	0.418	-0.475	0.259	0.067	-0.151	0.273	0.580
Protestant	0.090	0.454	0.842	0.156	0.191	0.413	-0.102	0.267	0.703	0.183	0.281	0.516
Philosophic background	−0.374	0.555	0.501	0.075	0.234	0.747	−0.742*	0.326	0.023	−0.308	0.344	0.371
Proportion SES neighbourhood students	—	—	—	—	—	—	—	—	—	—	—	—
Highest SES	0.007	0.008	0.372	0.007*	0.003	0.024	0.000	0.005	0.960	0.007	0.005	0.121
Lowest SES	−0.003	0.008	0.755	−0.005	0.003	0.155	−0.003	0.005	0.598	-0.008	0.005	0.116

Notes: *p < *0.05.* **p < *0.01.* ***p < *0.001.* [a]*ref one location,* [b]*ref 0–5%,* [c]*ref non-PDS,* [d]*ref no religious/philosophic background.*

teaching behavior. For each type of induction element, the average amount of perceived activities did not differ significantly between PDSs and non-PDSs, although for enculturation activities the differences were marginally significant, with PDSs offering somewhat less enculturation activities (b = −0.283, p = 0.052).

The bivariate regression models showed significant associations between some of the school characteristics and the number of perceived support activities (**Table 2**). First, schools with multiple locations offered less activities concerning workload reduction and enculturation than schools with a single location. Second, a larger proportion of BTs in the staff showed a negative association with the amount of perceived activities, in particular workload reduction and professional development plans. The binary models indicated a contrasting association of increasing average age of the male (negative) and female (positive) teaching staff with perceived activities provided to BTs, and also indicated a positive association between a greater proportion of female teachers and number of activities provided to BTs.

Catholic schools and schools with a specific educational philosophy provided less professional development plans and instruments compared with school without such background. The binary models indicated that the proportion of students living in the highest SES (positive) and lowest SES (negative) neighborhoods was inversely associated with the frequency of all induction elements. This means that the higher the SES of a school's student population, the more activities BTs perceive. However, nearly all associations were not significant in these binary models.

Predictive Value of School Characteristics for the Amount of Induction Activities

Estimates of the full regression models show that schools with multiple locations offer significantly less activities concerning workload reduction, enculturation and professional development plans and instruments. The larger the schools, the fewer of such activities are offered to BTs (**Table 3**). A larger proportion of BTs working at a school was associated with fewer perceived activities, with the strongest negative association with workload reduction and professional development plans. Increased average age of the male teaching staff was associated with fewer enculturation and professional development plans, whereas increased average age of female teachers was associated with more activities concerning support of effective teaching. A greater proportion of female teachers at a school was associated with more enculturation activities provided to BTs. Marginally significant positive associations between average age of female teachers and activities concerning enculturation (b = 0.052, p = 0.080), and professional development plans (b = 0.081, p = 0.054) were found.

Schools with a Catholic background or a specific educational philosophy offered less professional development plans and instruments than public schools. Finally, when accounting for the other school characteristics, no associations are found between the socioeconomic composition of schools' student population and the number of provided induction activities.

Determinants of provided induction activities for BTs working in a PDS and non-PDS context.

In order to understand whether the selected school characteristics had similar or different predictive values in different learning and support infrastructures, the full models were stratified for PDSs and non-PDSs. For PDSs, the full models showed that schools having multiple locations and at least 2,500 students offered significantly less activities concerning workload reduction and enculturation than those with a single location (**Table 4**). For non-PDSs having multiple locations but less than 2,500 students, this is the case for professional development plans as well (**Table 5**).

TABLE 3 | Linear regression models: reported support activities by BTs and school characteristics, four induction elements, full sample (N = 195).

	Workload reduction			Enculturation			Professionalization			Effective teaching		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***
Intercept	10.543*	4.242	0.013	5.636**	1.763	0.001	9.176***	2.493	0.000	3.437	2.763	0.213
School size[a]	—	—	—	—	—	—	—	—	—	—	—	—
Multiple locations <2,500 students	−0.769	0.430	0.074	−0.187	0.179	0.295	-0.388	0.253	0.125	0.028	0.280	0.921
Multiple locations>=2,500 students	−1.490***	0.404	0.000	−0.584***	0.168	0.001	−0.575*	0.237	0.015	−0.023	0.263	0.929
Proportion of beginning teachers (proxy)[b]	—	—	—	—	—	—	—	—	—	—	—	—
5–10%	−1.288***	0.416	0.002	−0.150	0.173	0.387	−0.687***	0.245	0.005	−0.619*	0.271	0.022
>10%	−2.012***	0.433	0.000	−0.479**	0.180	0.008	−1.057***	0.254	0.000	−0.649*	0.282	0.021
Age and gender composition teaching staff	—	—	—	—	—	—	—	—	—	—	—	—
Average age male teachers	−0.066	0.071	0.357	−0.099***	0.030	0.001	−0.140***	0.042	0.001	−0.059	0.047	0.205
Average age female teachers	0.058	0.072	0.415	0.052	0.030	0.080	0.081	0.042	0.054	0.095*	0.047	0.043
Proportion of female teachers (N)	0.037	0.032	0.241	0.028*	0.013	0.034	0.030	0.019	0.109	0.032	0.021	0.122
Professional development school[c]	—	—	—	—	—	—	—	—	—	—	—	—
Yes	0.054	0.330	0.870	−0.245	0.137	0.074	−0.055	0.194	0.777	−0.016	0.215	0.940
School identity[d]	—	—	—	—	—	—	—	—	—	—	—	—
Catholic	−0.684	0.417	0.101	−0.113	0.173	0.516	−0.533*	0.245	0.030	−0.222	0.272	0.415
Protestant	0.045	0.429	0.916	0.163	0.178	0.362	−0.098	0.252	0.698	0.195	0.279	0.485
Philosophic background	−0.562	0.535	0.293	0.076	0.222	0.734	−0.692*	0.314	0.028	−0.347	0.348	0.319
Proportion SES neighbourhood students	—	—	—	—	—	—	—	—	—	—	—	—
Highest SES	−0.001	0.009	0.921	−0.001	0.004	0.850	−0.010	0.005	0.068	0.001	0.006	0.916
Lowest SES	0.003	0.009	0.768	−0.002	0.004	0.572	−0.005	0.005	0.400	−0.006	0.006	0.277
Model summaries	—	—	—	—	—	—	—	—	—	—	—	—
LL (*df = 181*)	−430,82			−259,65			−327,17			−347,21		

Notes: *p < *0.05.* **p < *0.01.* ***p < *0.001.* [a]*ref one location,* [b]*ref 0–5%,* [c]*ref non-PDS,* [d]*ref no religious/philosophic background.*

TABLE 4 | Linear regression models: reported support activities by BTs and school characteristics, four induction elements, PDSs (N = 89).

	Workload reduction			Enculturation			Professionalization			Effective teaching		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***
Intercept	6.731	6.165	0.275	3.511	2.678	0.190	3.518	3.806	0.355	0.640	4.302	0.882
School size[a]	—	—	—	—	—	—	—	—	—	—	—	—
Multiple locations <2,500 students	−0.076	0.610	0900	0.438	0.265	0.098	0.530	0.377	0.160	0.565	0.426	0.185
Multiple locations>=2,500 students	−1.540**	0.496	0.002	−0.497*	0.215	0.021	−0.333	0.306	0.276	−0.105	0.346	0.760
Proportion of beginning teachers (proxy)[b]	—	—	—	—	—	—	—	—	—	—	—	—
5–10%	−0.901	0.551	0.102	−0.234	0.240	0.328	-0.666	0.340	0.050	−0.530	0.385	0.168
>10%	−1.421*	0.576	0.014	−0.283	0.250	0.258	-0.754*	0.356	0.034	−0.273	0.402	0.497
Age and gender composition teaching staff	—	—	—	—	—	—	—	—	—	—	—	—
Average age male teachers	0.190	0.113	0.092	−0.029	0.049	0.551	0.012	0.070	0.858	0.020	0.079	0.800
Average age female teachers	−0.127	0.114	0.267	0.000	0.050	0.999	0.036	0.070	0.608	0.087	0.080	0.274
Proportion of female teachers (N)	0.030	0.045	0.507	0.042*	0.020	0.032	0.032	0.028	0.254	0.027	0.031	0.385
School identity[c]	—	—	—	—	—	—	—	—	—	—	—	—
Catholic	−0.609	0.524	0.245	−0.190	0.228	0.403	−0.364	0.324	0.260	−0.177	0.366	0.629
Protestant	0.060	0.582	0.918	0.497	0.253	0.050	0.294	0.359	0.414	0.523	0.406	0.198
Philosophic background	−0.461	0.893	0.606	0.411	0.388	0.289	−0.495	0.551	0.369	−0.836	0.623	0.180
Proportion SES neighbourhood students	—	—	—	—	—	—	—	—	—	—	—	—
Highest SES	−0.007	0.015	0.643	0.003	0.006	0.637	−0.005	0.009	0.595	−0.008	0.010	0.436
Lowest SES	−0.001	0.014	0.949	−0.013*	0.006	0.029	−0.019*	0.008	0.027	−0.021*	0.009	0.029
Model summaries	—	—	—	—	—	—	—	—	—	—	—	—
LL (*df = 93*)	−184,668			−110,465			−141,753			−152,655		

Notes: *p < *0.05.* **p < *0.01.* ***p < *0.001.* [a]*ref one location,* [b]*ref 0–5%,* [c]*ref no religious/philosophic background.*

In schools where more than ten percent of the teaching staff were BTs, PDSs offered less activities concerning workload reduction and professional development plans (**Table 4**). This school characteristic had a stronger negative association with all provided activities in non-PDSs (**Table 5**). The average age of male and female teaching staff was not predictive in the amount of activities provided to BTs working in PDSs (**Table 4**), but did matter in a non-PDS context (**Table 5**). In non-PDSs, the amount of activities for enculturation, professional development plans and effective teaching was less with increasing age of male teachers, and more with increasing age of female teachers. In contrast, the positive association between a greater share of female teachers and enculturation activities appeared to be significant in PDS only (**Table 4**).

TABLE 5 | Linear regression models: reported support activities by BTs and school characteristics, four induction elements, non-PDSs (N = 106).

	Workload reduction			Enculturation			Professionalization			Effective teaching		
	Estimate	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***	***Estimate***	***SE***	***p***
Intercept	11.067	5.672	0.051	6.059**	2.137	0.005	11.051***	3.067	0.000	4.320	3.528	0.221
School size[a]	—	—	—	—	—	—	—	—	—	—	—	—
Multiple locations <2,500 students	−1.233*	0.580	0.034	−0.682*	0.218	0.002	−0.992**	0.314	0.002	−0.345	0.361	0.339
Multiple locations>=2,500 students	−1.311*	0.642	0.041	−0.564*	0.242	0.020	−0.714*	0.347	0.040	0.165	0.399	0.680
Proportion of beginning teachers (proxy)[b]	—	—	—	—	—	—	—	—	—	—	—	—
5–10%	−1.944***	0.598	0.001	−0.157	0.225	0.487	−0.920**	0.324	0.004	−0.921*	0.372	0.013
>10%	−2.349***	0.643	0.000	−0.647**	0.242	0.008	−1.401***	0.348	0.000	−0.847*	0.400	0.034
Age and gender composition teaching staff	—	—	—	—	—	—	—	—	—	—	—	—
Average age male teachers	−0.201*	0.091	0.027	−0.127***	0.034	0.000	−0.216***	0.049	0.000	−0.116*	0.056	0.040
Average age female teachers	0.179	0.091	0.050	0.082*	0.034	0.018	0.115*	0.049	0.020	0.126*	0.057	0.026
Proportion of female teachers (N)	0.052	0.045	0.248	0.025	0.017	0.137	0.041	0.024	0.093	0.039	0.028	0.159
Schoolculture[c]	—	—	—	—	—	—	—	—	—	—	—	—
Catholic	−0.730	0.624	0.242	−0.021	0.235	0.928	−0.663	0.338	0.050	−0.349	0.388	0.369
Protestant	−0.007	0.600	0.990	−0.148	0.226	0.511	−0.464	0.325	0.153	−0.161	0.373	0.665
Philosophic background	−0.708	0.682	0.299	−0.088	0.257	0.733	−0.825*	0.369	0.025	−0.372	0.424	0.380
Proportion SES neighbourhood students	—	—	—	—	—	—	—	—	—	—	—	—
Highest SES	0.007	0.012	0.545	-0.005	0.005	0.314	−0.013*	0.006	0.048	0.005	0.007	0.482
Lowest SES	0.010	0.012	0.411	0.004	0.005	0.360	0.006	0.007	0.330	0.004	0.008	0.601
Model summaries	—	—	—	—	—	—	—	—	—	—	—	—
LL (*df = 76*)	−237,05			−133,57			−171,89			−186,71		

Notes: *p < *0.05.* **p < *0.01.* ***p < *0.001.* [a]*ref one location,* [b]*ref 0–5%,* [c]*ref no religious/philosophic background.*

PDSs with a Protestant denomination provided significantly more activities for enculturation than PDSs with a Catholic or no religious/philosophic background (**Table 4**). Among non-PDSs, schools without a specific denomination or philosophy provided significant more activities for professional development plans compared with the remaining non-PDSs (**Table 5**).

Finally, the results showed that a higher proportion of students living in the lowest SES neighborhoods had a small negative association with the amount of activities concerning enculturation, professional development plans and instruments and classroom support provided to BTs working in PDSs (**Table 4**). Socioeconomic composition of students was not associated with perceived activities of BTs working in non-PDSs (**Table 5**).

DISCUSSION

With intensive support of regional teacher education institutes, 3-year induction programs were implemented in a large number of secondary schools (N = 447) in the Netherlands in order to improve teaching quality and increase retention of BTs. The elements of the arrangement were intended to be tailored to specific BT needs and the implementation were aimed to be adapted to each school's specific context in the design and implementation phase. As organisational contexts of schools are (to some extent) associated with particular demographic school characteristics that might facilitate or hinder the implementation of induction activities, the aim of this study was to approach the association of school factors with the implementation of induction arrangements as objectively as possible, by investigating the association between demographic characteristics of schools and provided support activities, as perceived by BTs themselves.

On average, participating schools offered at least half of the activities as prescribed in the four elements of the study's induction framework (RQ1). Findings indicate that increased school size decreases the provision of activities, which is in line with studies showing that teacher collaboration and engagement in common activities is less likely to occur in schools with more students (Fowler and Walberg, 1991; Lee et al., 1993; Maslowski, 2001; Opdenakker and Van Damme, 2007). Schools with more locations may have different local policies and different educational instructional design preferences which may cause less alignment in the induction arrangements across locations, leading to different perceptions of BTs concerning the available induction activities within one school (RQ2). As recommended by Kutsyuruba et al. (2016), strong school leadership can improve the implementation of induction arrangements, especially in larger schools having more locations. In addition, this study indicates that schools with a larger proportion of BTs provide less activities than schools with relatively few BTs. Providing support activities to a larger group of BTs requires a larger number of supportive colleagues and demands more detailed supportive infrastructures (RQ2). This may be relative easy to design on paper, but might be more difficult to put into practice, where more individual restrictions have to be accounted for, such as individual contract hours and teaching schedules. Guidelines adapted to this situation could assist school leaders in allocating sufficient induction activities for BTs.

The results also indicate that the teaching staff age composition matters. Older male teacher staff composition is negatively associated, and older female teacher staff composition is positively associated with BTs perceived activities (RQ2). This pattern could reflect that older female teachers more often take up supportive activities compared to older male teachers. As the teacher population is ageing in the Netherlands, with an increasing share of older male teachers in the near future, this finding implies that induction programmes may be less efficient in schools that have greater share of older male teachers. More in-depth investigations could provide more insight in and understanding of such mechanisms. The question whether schools encounter difficulties in meeting the specific support needs of both BTs and older male teachers as described by Day and Gu (2007), or whether interactions between older male teachers and BTs inhibit support activities as stated by Kardos et al. (2001), or whether other mechanisms hinder the implementation of induction arrangements in these contexts needs to be clarified in the future.

In this study, SES composition of students does not have a clear association with the type and amount of support activities provided to BTs (RQ2). Future studies should reveal the needs of BTs in more challenging school contexts, as it could be expected that BTs are in need for more support activities than currently provided.

In this study, differences in activities provided in PDSs and non-PDSs were specifically investigated, as the existing support and learning infrastructures were assumed to be different. It was argued that the existing infrastructure within PDSs to support student teachers would facilitate support activities for BTs, thereby making them less sensitive to school contextual factors compared to non-PDSs. This proved to be the case: for all four induction elements, the amount of activities provided was similar in both school contexts (RQ3). Within these two contexts, the amount of support provided was more strongly associated to school characteristics in non-PDS (RQ3). In non-PDSs, a higher share of BTs and a higher average age of the male teaching staff were negatively associated with perceived provided support activities (RQ3). For the PDSs the association of school characteristics with perceived provided support was smaller or non-existing, indicating that induction arrangements for BTs benefited from existing learning infrastructures aimed for student teachers. In the absence of such existing structures, policy makers and school leaders should be aware that particular school characteristics could hinder or enhance the implementation of support activities, which requires awareness and more intensive support. In PDSs, the socioeconomic composition of students appears to be associated with the support provision of BTs. For the benefit of education quality, it is relevant to gain more insight in the mechanism behind this finding, as schools with a more challenging student population are generally in greater need to provide adequate teacher support. Finally this study showed that PDSs are able to carry an additional responsibility (supporting BTs) besides the tasks and duties they already have in their collaboration with teacher education institutes to educating student teachers.

A major strength of this study is the large number of participating schools, enabling us to statistically adjust for various school characteristics in order to generate insights into differences between schools in providing induction arrangements to BTs. Another strength is that provided support is measured from the perspective of the BTs themselves, rather than from school leaders or formal documents. Having insight in how BTs perceive the provision of induction activities is relevant, as it generates information about the implementation process and effectiveness of such interventions.

A limitation of this study is that it can be questioned whether the reported support activities represent the induction arrangement of schools in an adequate manner, as not all BT's participating in the induction arrangement completed the support monitor. Potential unintended selective participation of BTs in this study could therefore lead to an over- or underreporting of actual provided induction activities, for instance when it were particularly those BT's that are positive about the induction arrangement who responded, or those who have a good relationship with the school environment. A second limitation of this study is that it did not capture BTs' perceptions of the quality and usefulness of the provided activities. Even though the support monitor contains questions about whether or not each offered induction activity was evaluated to be useful, this information could not be used in the analyses, as the vast majority of BTs reported the provided activities as being useful, resulting in too little variety in response patterns, for an aggregated school-level approach. Future research should apply an individual-level approach too, to provide insight in individual determinants of differences in support provision, where the perceived quality of provided induction activities is incorporated.

The insights gained from this study are relevant for mapping school differences in induction arrangements and can therefore be informative for tailoring induction arrangements to particular school contexts in order to assure optimal development of teaching skills of all BTs.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the respondent's were not asked for permission to share data. Requests to access the data should be directed to m.helms-lorenz@ rug.nl.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics commission Teacher Education department.

The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MV was in the lead in the writing process and the analyses. MH-L shares the intellectual ownership and helped shape the manuscript, and is the project leader responsible for the research proposal, the data collection and its national execution.

FUNDING

This project had been fully funded by the Ministry of Education. Codes OND/ODB-2013/45916 U (804A0-45835) and OND/ODB-13/19888 U (804A0-42807).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2021.744347/full#supplementary-material

REFERENCES

Birkeland, S., and Feiman-Nemser, S. (2012). Helping School Leaders Help New Teachers: A Tool for Transforming School-Based Induction. *The New Educator.* 8 (2), 109–138. doi:10.1080/1547688X.2012.670567

Bonesrønning, H., Falch, T., and Strøm, B. (2005). Teacher Sorting, Teacher Quality, and Student Composition. *Eur.Econ. Rev.* 49, 457–483. doi:10.1016/S0014-2921(03)00052-7

Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2005). Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students. *Am. Econ. Rev.* 95 (2), 166–171. doi:10.1257/000282805774669628

Bryk, A., Camburn, E., and Louis, K. S. (1999). Professional Community in Chicago Elementary Schools: Facilitating Factors and Organizational Consequences. *Educ. Adm. Q.* 35 (5), 751–781. doi:10.1177/0013161X99355004

Centraal Bureau voor de Statistiek (2018). *Werkzame Beroepsbevolking Beroep, Persoonskenmerken, Kenmerken Werkkring* [Employed Labourforce, Occupation, Individual Characteristics, Work Characteristics][Werkzame Beroepsbevolking, Beroep]. Available at: https://opendata.cbs.nl/statline/portal.html?_la=nl&_catalog=CBS&tableId=82808NED&_theme=14 (Accessed December 2018)

Creemers, B., and Kyriakides, L. (2010). School Factors Explaining Achievement on Cognitive and Affective Outcomes: Establishing a Dynamic Model of Educational Effectiveness. *Scand. J. Educ. Res.* 54 (3), 263–294. doi:10.1080/00313831003764529

Danhier, J. (2016). Teachers in Schools with Low Socioeconomic Composition: Are They Really that Different? *Eur. Educ.* 48 (4), 274–293. doi:10.1080/10564934.2016.1248231

Day, C., and Gu, Q. (2007). Variations in the Conditions for Teachers' Professional Learning and Development: Sustaining Commitment and Effectiveness over a Career. *Oxford Rev. Educ.* 33 (4), 423–443. doi:10.1080/03054980701450746

Dienst Uitvoering Onderwijs (2016). Alle Vestigingen VO [School Locations Secondary Education][*Adressen Voortgezet Onderwijs*]. Available at: https://duo.nl/open_onderwijsdata/databestanden/vo/adressen/ (Accessed June 2017).

Dienst Uitvoering Onderwijs (2015a). Leerlingen per vestiging naar postcode en leerjaar - 2014-2015 [Students per location by postcode and year][*Leerlingen in het voortgezet onderwijs*]. Available at: https://duo.nl/open_onderwijsdata/databestanden/vo/leerlingen/ (Accessed June 2017).

Dienst Uitvoering Onderwijs (2018). Onderwijspersoneel in Aantal Personen [Figures Staff Per School Location Secondary Education][*Onderwijspersoneel In Aantal Personen*]. Available at: https://duo.nl/open_onderwijsdata/databestanden/vo/onderwijspersoneel/ (Accessed January 2018).

Dienst Uitvoering Onderwijs (2015b). Prognose Aantal Leerlingen in Het Voortgezet Onderwijs [Projections Number of Students in Secondary Education][*Leerlingen In Het Voortgezet Onderwijs*]. Available at: https://duo.nl/open_onderwijsdata/databestanden/vo/leerlingen/ (Accessed June 2017).

Firestone, W. A., and Louis, K. S. (1999). "Schools as Cultures," in *Handbook of Research on Educational Administration*. Editors J. Murphy and K. S. Louis. 2nd ed. (San Francisco: Jossey-Bass), 297–322.

Fowler, W. J., and Walberg, H. J. (1991). School Size, Characteristics, and Outcomes. *Educ. Eval. Pol. Anal.* 13 (2), 189–202. doi:10.3102/01623737013002189

Hanushek, E. A., Kain, J. F., and Rivkin, S. G. (2004). Why Public Schools Lose Teachers. *J. Hum. Resour.* XXXIX (2), 326–354. doi:10.3368/jhr.XXXIX.2.326

Harmsen, R., Helms-Lorenz, M., Maulana, R., and Veen, K. (2018). The Longitudinal Effects of Induction on Beginning Teachers' Stress. *Br. J. Educ. Psychol.* 89, 259–287. doi:10.1111/bjep.12238

Helms-Lorenz, M., Buitink, J., t' Hart, W., van Mourik, C. A., Schuurman, G., and Ewalts, H. (2015). *Begeleiding van beginnende docenten, een kader voor de ontwikkeling van school-specifieke inductiearrangementen* [Supporting beginning teachers, a framework for the development of school specific induction arrangements]Groningen: Lerarenopleiding Rijksuniversiteit Groningen

Helms-Lorenz, M., Koffijberg, J., Schellings, G., van der Wolk, W., Aarts, R., Runhaar, P., et al. (2018a). "De Driejarige Begeleiding Van Startende Leraren in Nederland [Supporting Beginning Teachers in the First Three Years of Their Professional Practice in the Netherlands]," in Start To Teach: Inspiratiegids Over Aanvangsbegeleiding in Het Onderwijs *[Start to Teach: An Inspirationguide for Induction Activities in Education]*. Editors S. De Vos, J. de Wilde, and S. Beausaert (Antwerpen-Apeldoorn: Garant Publishers), 55–77.

Helms-Lorenz, M., Slof, B., Vermue, C. E., and Canrinus, E. T. (2012). Beginning Teachers' Self-Efficacy and Stress and the Supposed Effects of Induction Arrangements. *Educ. Stud.* 38 (2), 189–207. doi:10.1080/03055698.2011.598679

Helms-Lorenz, M., van de Grift, W., Canrinus, E., Maulana, R., and van Veen, K. (2018b). Evaluation of the Behavioral and Affective Outcomes of Novice Teachers Working in Professional Development Schools versus Non-professional Development Schools. *Stud. Educ. Eval.* 56, 8–20. doi:10.1016/j.stueduc.2017.10.006

Helms-Lorenz, M., van de Grift, W., and Maulana, R. (2016). Longitudinal Effects of Induction on Teaching Skills and Attrition Rates of Beginning Teachers. *Sch. Effectiveness Sch. Improvement.* 27 (2), 178–204. doi:10.1080/09243453.2015.1035731

Ingersoll, R. M., and Smith, T. M. (2004). Do teacher Induction and Mentoring Matter? *NASSP Bull.* 88 (638), 28–40. doi:10.1177/019263650408863803

Ingersoll, R. M., and Strong, M. (2011). The Impact of Induction and Mentoring Programs for Beginning Teachers. *Rev. Educ. Res.* 81 (2), 201–233. doi:10.3102/0034654311403323

Inspectie van het Onderwijs (2017). De Staat Van Het Onderwijs *[The State of Education in the Netherlands]*. Utrecht: Ministerie van Onderwijs, Cultuur en Wetenschap.

Johnson, S. M., Kardos, S. M., Kauffman, D., Liu, E., and Donaldson, M. E. (2004). The Support gap: New Teacher's Early Experiences in High Income and Low Income Schools. *Educ. Pol. Analyses Arch.* 12, 61. doi:10.14507/epaa.v12n61.2004

Kardos, S. M., Johnson, S. M., Peske, H. G., Kauffman, D., and Liu, E. (2001). Counting on Colleagues: New Teachers Encounter the Professional Cultures of Their Schools. *Educ. Adm. Q.* 37 (2), 250–290. doi:10.1177/00131610121969316

Kessels, C. (2010). The Influence of Induction Programs on Beginning Teachers 'Well-Being and Professional Development. Universiteit Leiden. Available at: http://hdl.handle.net/1887/15750.

Kraft, M. A., and Papay, J. P. (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to

Teaching Experience. *Educ. Eval. Pol. Anal* 36 (4), 476–500. doi:10.3102/0162373713519496

Kutsyuruba, B., Godden, L., Covell, L., Matheson, I., and Walker, K. (2016). *An International Systematic Review of Research; Understanding the Contextual Factors within Teacher Induction and Mentoring Programs*. Kinston and Saskatoon: Queen's University and University of Saskatchewan.

Lankford, H., Loeb, S., and Wyckoff, J. (2002). Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis. *Educ. Eval. Pol. Anal.* 24 (1), 37–62. doi:10.3102/01623737024001037

Lee, V. E., Bryk, A. S., and Smith, J. B. (1993). Chapter 5: The Organization of Effective Secondary Schools. *Rev. Res. Educ.* 19, 171–267. doi:10.3102/0091732x019001171

Loeb, S., Darling-Hammond, L., and Luczak, J. (2005). How Teaching Conditions Predict Teacher Turnover in california Schools. *Peabody J. Educ.* 80 (3), 44–70. doi:10.1207/s15327930pje8003_4

Maslowski, R. (2001). *School Culture and School Performance*. (Doctoral Dissertation). Available at: https://ris.utwente.nl/ws/portalfiles/portal/6074941

Maulana, R., Helms-Lorenz, M., and van de Grift, W. (2015). A Longitudinal Study of Induction on the Acceleration of Growth in Teaching Quality of Beginning Teachers through the Eyes of Their Students. *Teach. Teach. Educ.* 51, 225–245. doi:10.1016/j.tate.2015.07.003

Muijs, D., Harris, A., Chapman, C., Stoll, L., and Russ, J. (2004). Improving Schools in Socioeconomically Disadvantaged Areas ? A Review of Research Evidence. *Sch. Effectiveness Sch. Improvement* 15 (2), 149–175. doi:10.1076/sesi.15.2.149.30433

Nederlands-Vlaamse Accreditatieorganisatie (2009). Toetsingskader Opleidingsschool *[Accreditation Framework for Professional Development School]*. Den Haag Netherlands: NVAO, Afdeling Nederland.

Opdenakker, M.-C., and Damme, J. (2001). Relationship between School Composition and Characteristics of School Process and Their Effect on Mathematics Achievement. *Br. Educ. Res. J.* 27 (4), 407–432. doi:10.1080/01411920120071434

Opdenakker, M. C., and Damme, J. V. (2007). Do school Context, Student Composition and School Leadership Affect School Practice and Outcomes in Secondary Education? *Br. Educ. Res. J.* 33 (2), 179–206. doi:10.1080/01411920701208233

Organisation for Economic Co-operation and Development (2018). Effective Teacher Policies: Insights from PISA. PISA, OECD Publishing. doi:10.1787/9789264301603-en

Organisation for Economic Co-operation and Development (2016). Netherlands 2016: Foundations for the Future, Reviews of National Policies for Education. Paris: OECD Publishing. doi:10.1787/9789264257658-en

Richardson, R. C., Glessner, L., and Tolson, H. (2007). Stopping the Leak: Retaining Beginning Teachers. *Ajte* 32 (2). doi:10.14221/ajte.2007v32n2.2

Ronfeldt, M., Kwok, A., and Reininger, M. (2016). Teachers' Preferences to Teach Underserved Students. *Urban Educ.* 51 (9), 995–1030. doi:10.1177/0042085914553676

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., and Feng, L. (2012). Value Added of Teachers in High-Poverty Schools and Lower Poverty Schools. *J. Urban Econ.* 72, 104–122. doi:10.1016/j.jue.2012.04.004

Scafidi, B., Sjoquist, D. L., and Stinebrickner, T. R. (2007). Race, Poverty, and Teacher Mobility. *Econ. Educ. Rev.* 26, 145–159. doi:10.1016/j.econedurev.2005.08.006

Scheerens, J. (2016). *Educational Effectiveness and Ineffectiveness. A Critical Review of the Knowledge Base*. Netherlands: Springer.

Sociaal en Cultureel Planbureau (2014). Statusscores, Peildatum 2014 [*Statusscores 2014*]. Available at: https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores (Accessed November 2017).

Sykes, B., and Kuyper, H. (2013). School Segregation and the Secondary-School Achievements of Youth in the netherlands. *J. Ethnic Migration Stud.* 39 (10), 1699–1716. doi:10.1080/1369183X.2013.833707

Sykes, B., and Musterd, S. (2011). Examining Neighbourhood and School Effects Simultaneously. *Urban Stud.* 48 (7), 1307–1331. doi:10.1177/0042098010371393

Van den Berg, D., Defourny, R., Kuipers, Z., and Stevenson, S. (2015). *Regionale Arbeidsmarktrapportages Voortgezet Onderwijs 2015 [*Projections Regional Labourmarket Secondary Education 2015. Heerlen: VOION

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Education

ORIGINAL RESEARCH
published: 19 October 2021
doi: 10.3389/feduc.2021.737163

Counselor-Perceived Teacher Actions Needed to Carry out Educational and Vocational Guidance in Secondary Schools in Spain

Carolina Falcón-Linares, Ana Rodríguez-Martínez, Alejandra Cortés-Pascual and Alberto Quilez-Robres*

Department of Educational Sciences, Faculty of Education, University of Zaragoza, Zaragoza, Spain

After analyzing the scientific literature referring to guiding actions developed in secondary education, there is an appreciable lack of research on process efficiency performed by teachers. This study explores the practice of tutoring, which in Spain is coordinated by counselors but the responsibility of teachers. An *ad hoc* semistructured questionnaire was designed, validated and applied within the framework of a collaboration agreement between professional institutions of counselors and researchers. The questionnaire was answered by 102 counselors who are heads of guidance departments because the insight of counseling experts was desired. The quantitative and qualitative data analyses indicated correlations between the profile of the counselors and the performance of tutoring, described the content and predominant intervention models and identified needs for improvement. A much more intensive and direct intervention by counselors is necessary in a system in which teachers are not trained enough and time dedicated to tutorial action is deficient. Notably, counselors coordinate and supervise tutoring, but they feel overwhelmed by the lack of concrete norms and their workload.

Keywords: school counselor, school tutoring, vocational guidance, secondary education, qualitative research

OPEN ACCESS

Edited by:
Ridwan Maulana,
University of Groningen, Netherlands

Reviewed by:
Serafina Pastore,
University of Bari Aldo Moro, Italy
Seyeoung Chun,
Chungnam National University, South Korea

****Correspondence:***
Carolina Falcón-Linares
cfalcon@unizar.es

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *06 July 2021*
Accepted: *29 September 2021*
Published: *19 October 2021*

Citation:
Falcón-Linares C, Rodríguez-Martínez A, Cortés-Pascual A and Quilez-Robres A (2021) Counselor-Perceived Teacher Actions Needed to Carry out Educational and Vocational Guidance in Secondary Schools in Spain. Front. Educ. 6:737163. doi: 10.3389/feduc.2021.737163

INTRODUCTION

Co-ordinated tutoring is one of the most outstanding functions of teachers. Thus, it constitutes a fundamental support of teaching action from the point of view of professional self-efficacy, teacher-student relations and a positive school culture (OECD, 2009). The purpose of this study is to determine, in an exploratory sense, how tutoring in compulsory secondary education is developed from the perspective of counselors. The research is qualitative, conducted in Spain and involves counselors from almost the entire national territory. The starting point is a set of research questions (RQ) that define the research problem and guide the investigative process: RQ1) Are there relationships between counselor characteristics and tutoring practices? RQ2) What tutoring models predominate in relation to content, processes and work strategies? RQ3) What are the needs of the teachers as tutors?

Guidance and tutorial action are topics pending reflection, research, revaluation and updating to improve practices. Currently, these activities are incorporated into secondary education as a model of school integration in almost all European countries, although the pace of implementation varies (Hearne and Neary, 2021). Especially relevant is the integration of vocational guidance at this stage from a comprehensive perspective, and yet, it continues to be considered an excessively politicized

process (Sultana, 2014). Tutorial action is a first-level guidance intervention, carried out by teachers under the coordination of school counselors. Tutoring is offered to students and their families. It is mainly performed by a teacher-tutor but involves all teachers. To improve tutoring processes, it is necessary to start with a contextualized needs analysis and European strategic recommendations on life-long guidance (González-Benito and Vélaz de Medrano, 2014; European Commission, 2020).

Counselors, as those responsible for and the coordinators of tutorial actions, demand the clarification of and improvements in tutoring functions, competencies, strategies and organizational structures. Over the years, not only has the educational scenario change substantially, but research has continually contributed new theoretical and strategic points of view on guidance (Haug and Plant, 2016; Perry and Shannon, 2017). Priorities in vocational guidance, inclusive education and assessment procedures, among others, are evolving (ETF, 2020; Council of the European Union, 2021). However, it is not known whether this theoretical-strategic progress is transforming the efficacy of teachers in their role as tutorsin the Spanish context.

The main beneficiaries of counseling, that is, adolescents, have specific needs that go beyond what is related to their professional career. They need to feel that they belong to a safe context, that they have opportunities and new challenges and that they do not see their leisure options limited and feel they can overcome cultural and socioeconomic prejudices (Aguiar and Conceição, 2015). In addition, the need for guidance within the school environment is increased in students with low academic performance, less parental support and low family income (Park et al., 2018). A change in focus is urgently needed for vocational guidance: from counseling individuals regarding important decisions to helping them develop skills to understand and make decisions autonomously and with resilience. This new paradigm guarantees positive youth development, proactivity, self-determination, entrepreneurship, self-efficacy and motivation (Andrei and Solberg, 2019). In fact, being career ready requires acquiring employability skills associated with one's career goals (Castellano, Richardson, Sundell and Stone III, 2017).

TUTORIAL ACTION IN SECONDARY EDUCATION

During the second half of the last century, secondary education emerged linked to higher education. The preparatory purpose of this stage and its structure in cycles, modalities or elective pathways condition the functions of teachers. The criticisms of the poor pedagogical training of teachers in inclusion, equity and sustainability have a long history and continue today (Lorenzo et al., 2015; Silva et al., 2017). How teachers are trained has drawn much controversy and has been complicated by broad international heterogeneity regarding theoretical meaningfulness and practical organization. The most substantial difference refers to the moment of teacher selection, which can be before, during or after acquiring training (Ries et al., 2016), which determines the importance given to psychopedagogical knowledge as a basis or as a complement.

Some studies seek to understand how tutoring is valued, in terms of its meaningfulness and importance, by the main protagonists: tutors and students. These studies aim to determine the time dedicated to guidance and the difficulties encountered by tutors when tutoring. In general, there is widespread agreement on the importance of tutoring for the proper functioning of schools and for the quality of teaching. Moreover, the job satisfaction and the teachers' belief in their own effectiveness are fairly similar across countries (OECD, 2009). The most common problems are the lack of motivation of some teachers who do not want to be tutors because of the responsibility entailed, a lack of recognition by teachers and insufficient time allocated for tutorial work (López, 2013). Additionally, the gap between the relevance that students attach to vocational guidance during high school and the low satisfaction with the work done by their tutors and school counselors is substantial (Rodríguez et al., 2018). On the other hand, professional counselors state that adolescents who develop and participate in personal and teacher-led projects acquire transversal skills that improve their level of employability (Rodríguez et al., 2021).

Vélaz-de-Medrano et al. (2018) evaluate the performance of tutorial action based on data provided by tutors. The authors use an *ad hoc* methodology to determine the level of tutoring performance based on tutor perceptions. The study has a large sample of compulsory secondary education teachers in the Valencian Community (Spain). The perception of the tutors reveals a mismatch between the real and ideal dedication to personalizing teaching to achieve the maximum comprehensive development of the students. The study notes various causes that are related in a circular, nonlinear way: high levels of self-demand, insufficient training, lack of time and complex contexts. Teachers see themselves as more entrepreneurial, collaborative and proactive in relation to their own professional development when they formulate learning goals for themselves and the culture promotes transformational leadership for teacher learning (Louws et al., 2017). However, several demographic variables have had varying degrees of influence on teachers' occupational interests in secondary schools. Teachers differed significantly in their realistic, investigative, artistic, enterprising and occupational interests on the basis of gender and rank (Oyifioda and Iornenge, 2020).

In relation to family engagement, participation models are more complex in secondary school than in previous stages, increasing diversity among school contexts. The relationship between family participation and variables such as school climate, context and academic performance also increases. Improving the relationships between counselors, guardians and families is associated with having received training in skills to provide services to families and expand parents' expectations about counseling (Castro et al., 2015; Vaishnavi and Kumar, 2018).

Several studies report the prevalence of a more academic needs-focused guidance model rather than personal and interpersonal development. Students perceive that the work of

a counselor is aimed at enabling them to face challenges that may appear throughout their academic journey. The view of the counselor as a school psychologist is maintained in many institutions, where they provide personal and emotional adjustment and are only sought for advice to resolve conflicts and obtain help (Amor and Serrano, 2020; Auger et al., 2018; Carey et al., 2018). Many teachers also perceive counselors this way in educational institutions, demonstrating that much remains to be done regarding the review, definition and clarification of the functions and competencies of counselors (Johnson et al., 2010). When tutoring is analyzed from the perspective of counselors, it is observed that they dedicate a large part of their time to carrying out tasks related to the collection, synthesis, transformation and dissemination of resources to teachers. Counselors define themselves as sources of information and knowledge but also as disseminators of ideas, curriculum materials and research findings. They perform professional consulting functions that facilitate access to the appropriate tools for teachers to work autonomously and efficiently. However, in addition, they believe that there are high demands on them regarding activities outside the scope of their responsibilities and that spending time on these activities has a negative impact on their efficiency as counselors (Benigno, 2017; Hernández and Mederos, 2018). A survey of school counselors analyzed self-efficacy with school climate, counselor roles, and a variety of demographic variables. Results indicated that outcome expectancy for counselor behavior was predicted by a high degree of support from school staff and administrators and fewer nonrelated counseling activities performed by school counselors (Sutton and Fall, 1995).

With regard to guidance in secondary education, professional counseling deserves special attention. Currently, it is conceptualized, at least in the theoretical field, with a proactive, procedural, mediating and accompanying character in all academic stages. It is also necessary to include education in work values from a reflective and non-determining perspective. It is about preventing vocational changes and the abandonment of university careers (Cano, Orejudo and Cortés, 2019). Stable and lasting positive effects have been found in interventions that introduce professional counseling with group and individual activities into school hours (Andrei and Solberg, 2019; Babarović et al., 2020). However, more varied technical resources are needed, such as interviews, talks, meetings with families, intervention programs with TIC resources and with other conventional materials (Rodríguez et al., 2018; Costa et al., 2020).

Some current international experiences show the benefits of integrating guidance in teaching and the curriculum from an early age. In an educational institution in Wisconsin, an evaluative investigation of a program that integrates guidance beginning in elementary school was developed. It is a community plan aimed at working individually on academic, personal, social and career development. The results are very positive and contribute activities and resources to the international counseling community. Among its conclusions is the importance of cooperation between all teachers and counselors as well as the importance of training and the inclusion of teachers and families throughout the process (Boland et al., 2019). Another example is a publication by Sultana R, (2018) that represents a complete framework for career guidance in middle and high schools. It provides six fields of action for the development and improvement of school counseling services, with recommendations and resources for teachers and those responsible for education. As a transversal axis, reflective practice is emphasized to ensure quality in any guidance program.

When all teachers are involved in counseling, the principles of development and prevention are met, but specific training and time are needed to achieve this. It is essential to integrate educational programs with the curriculum, under the coordination of counselors, to ensure the evaluation and acquisition of life skills (Chan et al., 2011; Jariot, 2010). The training of secondary school teachers must include greater pedagogical and counseling knowledge instead of focusing almost exclusively on the didactics of a subject. The change in approach is needed from which the role and basic functions of teachers are contemplated, as well as the integration of content, content pedagogy and pedagogical knowledge (Imbernón, 2019; Neumann et al., 2019).

As seen in the theoretical framework, both the content and the model of action from which tutoring is addressed are in full practical and investigative development. To this reality is added the diversity of contexts that characterize secondary education in our country. In this study, three specific research objectives are proposed: (a) carry out an exploratory analysis of the guidance action developed by secondary education teachers in Spain in relation to the current theoretical framework; (b) describe the types of actions, strategies and instruments most used by teachers when tutoring, as well as emerging needs; and (c) analyze the relationship between the vision of guidance based on counselors, teachers and educational policy.

METHODS

The design places the counselor as the center of the study to serve a dual purpose: first, to obtain the professional perspective of counselors regarding tutorial actions, as they direct and coordinate tutorial action as guidance specialists; and second, to complete the existing research framework on tutorial action from other perspectives: teachers, students and families. An exploratory and descriptive study is conducted using an *ad hoc* questionnaire aimed at counselors who are heads of the guidance department of secondary schools. The educational context is the stage that includes students between 12 and 16 years of age; this context serves as compulsory education and is preparatory to the labor market and postcompulsory training.

The study began when the researchers requested professional support from two reference entities in counseling in our country: the local Asociación Aragonesa de Psicopedagogía (AAPS) (Aragonese Association of Psychopedagogy) and the national Confederación de Organizaciones de Psicopedagogía y Orientación de España (COPOE) (Confederation of Organizations of Psychopedagogy and Counseling of Spain). A

collaboration agreement was signed between the research team of the university and these organizations. As a result of this relationship, research was developed, responding to common interests, goals and intervention routes, and the questionnaire was validated and distributed throughout Spain.

Sample

The respondants were 102 secondary education counselors in Spain (65 women and 37 men). Seventy-six percent currently exercise their professional functions as heads of guidance departments; twenty percent have held this position in the last 5 years, and the remaining 6% did so previously. Forty-six percent of the sample had more than 21 years of experience in counseling. The questionnaire was disseminated through the Spanish Psychopedagogy Association network; therefore, there were participants from 13 autonomous communities. Of the counselors, 88% work in a public institution, 12% in a subsidized institution and 2% in a private institution. The response to the questionnaire was voluntary with selfless collaboration. The professional experience of the informants and their representativeness of the group of secondary school counselors in the national territory in various types of educational centers was evident.

Instrument

An *ad hoc* questionnaire was developed that specifically addressed the purposes of this research. First, a deconstructive reflection was generated based on the guidance actions that correspond to teachers, in relation to the three specific objectives, obtaining several investigative nuclei. These nuclei of interest are: a) tutoring content, b) the types of actions aimed at students and families, c) the sources and instruments to obtain information, d) the effectiveness of the processes (evaluation, resources, training and regulations), and e) detecting needs.

After the preliminary design of the questionnaire by the research team, it was subjected to evaluation by guidance experts from various fields who were intentionally selected. Some of them performed the review and provided contributions individually through email. The others participated in a focus group in which the nuclei of interest and the items on the instrument were discussed. The group of experts comprised 12 professionals: 3 professors from the University of Saragossa with a research career in counseling and 9 secondary education counselors. Of the latter, 5 worked in the public institutions, 2 in subsidized schools and 2 in private nonsubsidized institutions. A semi-structured panel of experts was carried out with several phases: I) Training session on the objectives of the research. II) Presentation of the preliminary questionnaire with the items associated with the nuclei of interest. III) Dynamics of participation and discussion on the validity of the items, their content, writing and organization within the questionnaire.

After the evaluation of the questionnaire by the experts, the final version was constructed by the research group and submitted to a final review by the presidents of the collaborating entities (AAPS and COPOE). The complete process of construction and validation of the questionnaire was carried out between April and June 2019. The collaborators disseminated the instrument for data collection through an online form between October 2019 and February 2020.

The questionnaire is semistructured and consists of two parts. The first collects demographic and professional data of the participants. The second contains 13 statements that include 49 items with different response formats. There are closed-ended choice items that condition a response (polytomous and dichotomous), other items that collect frequency data, graduated rating scales and, finally, open-ended questions.

The validity of the instrument is based on triangulation as the most supported reliability strategy in qualitative methodology. The contributions of the theoretical framework, the observations of the experts and the methodological combination that integrates quantitative and qualitative data in the study of the same object have been contrasted (Flick, 2009; Aguilar and Barroso, 2015).

Data Analysis

Taking into account the diversity of response types to the questionnaire items, several types of analysis were performed. Regarding the first research question (Are there relationships between counselor characteristics and tutoring practices?) quantitative analysis, without predetermined hypotheses and of an exploratory nature, was used to reveal emerging links between participant data and tutorial actions. This analysis was performed using the statistical software IBM SPSS Statistics 25.0. In relation to the second research question (What tutoring models predominate in relation to content, processes and work strategies?) a frequency analysis was carried out that provided percentages and facilitated the comparison between the categories proposed for each item. Third, the last research question (What are the needs of the teachers as tutors?) was answered by analyzing the data provided in the open-ended questions. According to the nature of the responses with narrative data, the qualitative analysis was analytical and inductive by taking Grounded Theory (Strauss, 1987; Strauss and Corbin, 1990) as method. Therefore, work began by using the key themes supported by the experts in the questionnaire validation -skills requirements, resources, evaluation, commitment, support and acknowledgement-. Then, possible connections among the emerging categories were explored until a two-level coding. The generated map was verified in a last round of data review; and categories were constantly refitted. The entire process was developed and revised by research authors.

RESULTS

In relation to RQ1, a first correlation analysis (**Table 1**) yielded data that link certain descriptive aspects of the counseling profile with the transversal topics addressed by the tutor. The gender of the counselor is related to issues such as nutrition and sports (ρ = 0.198; $p < 0.05$). Another relevant aspect in the counseling function, experience, is inversely associated with addiction prevention intervention ($\rho = -0.200$; $p < 0.05$). There is a

TABLE 1 | *Spearman ρ* (rho) Correlations between Counselor Profile and the Transversal Content Addressed by Tutors.

	Age	Gender	Experience	Degree	Teaching	Innovation projects
Emotional and Social Skills	–0.061	–0.009	–0.031	–0.076	0.043	0.030
Coexistence	0.105	0.044	0.029	0.052	0.104	0.046
Values	–0.076	–0.060	–0.091	0.224[a]	0.099	0.119
Addiction Prevention	–0.183	0.056	–0.200[a]	0.138	–0.026	**–0.250**[a]
Affective-Sexual	–0.076	0.078	–0.035	0.127	–0.103	–0.137
Gender Equality and Violence	0.154	–0.024	0.194	–0.025	0.072	0.068
Nutrition and Sports	0.141	0.198[a]	0.007	0.129	0.053	–0.015
Self-knowledge and decisions	0.055	–0.014	0.029	-–0.208[a]	–0.007	0.156
Communication	0.063	0.053	0.074	–0.085	**–0.228**[a]	–0.027
Academic Pathways and Options	0.026	–0.032	0.096	0.041	0.181	–0.076
Reinforcement and Study Strategies	0.127	–0.110	0.066	0.047	–0.114	–0.020

[a]p < *.05.Bold values represents the explanation of results.*

positive relationship between the academic degree attained by counselors and the development of values ($\rho = 0.224$; $p < 0.05$) and an inverse relationship with self-knowledge and decision-making ($\rho = -0.208$; $p < 0.05$). Counselor teaching has an inverse relationship with the development of communication skills ($\rho = -0.228$; $p < 0.05$). Finally, participation in innovation projects is also inversely related to addiction prevention ($\rho = -0.250$; $p < 0.05$).

However, to determine the direction of these relationships and confirm the existence or absence of significant differences, a second statistical analysis of variance (ANOVA) was performed. A significant difference ($F = 4.101$; $p < 0.05$) in the relation between counselor gender and the introduction of transversal topics, such as nutrition and sports, was confirmed. In this case, the majority of female counselors introduce this topic. On the other hand, the relationship between counseling experience and preventive treatment of addiction was confirmed, with the existence of significant differences ($F = 2.215$; $p < 0.05$): the less experience a counselor had (less than 5 years), the greater the dedication to these topics. However, the relationship between the counselor's degree and the development of values and the activities related to self-knowledge and decision-making was not confirmed, as no significant differences were found ($F = 2.191$; $p = 0.094$ and $F = 1989$; $p < 0.121$). Although there was a positive relationship for teaching degree based on Spearman's statistic, this result was not confirmed in the ANOVA. On the other hand, the relationship between nonteaching and the development of topics focused on communication was confirmed. A significant difference was found ($F = 5.432$; $p < 0.05$) for those who do not teach being more dedicated to topics of communication, as they are the ones who convey the need to work transversally. Finally, the relationship between participation in innovation projects and dealing with topics, such as the prevention of addiction, was not confirmed, and no significant differences were found between those who did and those who did not participate in this type of project ($F = 0.015$; $p = 0.902$).

In the following block of results RQ2 is answered. The first three nuclei of interest are described with respect to the tutorial action developed by the teachers, from the perspective of the counselors. Tutoring content, actions aimed at students and families, and the sources and instruments used by teachers to better understand their students are considered.

In relation to the content worked on by the teachers responsible for tutoring, they mostly focused on coexistence (87.3%). Emotional and social skills education also carried significant weight (62.7%). More than half of the participating counselors affirmed that the action of tutors is aimed at academic and professional pathways (58.8%) and carrying out preventive actions related to gender violence (52.9%). However, the other content considered relevant in the theoretical framework of tutorial actions were not significantly reflected in the practices performed by tutors, as follows, in decreasing order: academic reinforcement and training in study strategies (45.1%), affective-sexual education (38.2%), values education (36.3%), addiction prevention (34, 3%), guidance for self-knowledge and decision-making (33.3%), education in nutrition and sports (6.9%), training in communication skills (8.8%) and development of entrepreneurial attitude (0%).

The same items were studied in regard to the dedication of teachers who do not have to tutor. Overall, less than half of counselors feel that teachers exert some type of guidance. The only area to which more time and effort was dedicated was coexistence, mentioned by 58.8% of the counselors in relation to nontutoring teachers. Time dedicated to training activities in the emotional and social fields was below 33%. However, education in equality and gender violence prevention as well as educational reinforcement actions were closer to those performed by the tutors (44.1 and 40.2%, respectively). The other content decreased proportionally to the above with respect to tutors. In general, terms, a large majority of teachers do not exercise counseling functions with their students, reducing their teaching performance.

This thematic content was not addressed with the same dedication of time and resources in the four courses that were studied. As the most relevant synthesis, the three thematic blocks prioritized in each academic year were arranged in order from highest to lowest. In the first year, 67.64% prioritized coexistence and social skills education activities. The next topic in terms of dedication was training in learning strategies, with 17.64%, and the third was adaptation to the institute and the educational stage, with 7.84%. In the second year, the three thematic blocks in order of priority were coexistence and education in social skills (42.15%), addiction prevention (13.72%) and affective-sexual education (12.74%). In the third year, there were three blocks that shared a leading role: coexistence and education in social

TABLE 2 | Frequencies obtained for the different types of intervention with students.

	Not at all	Slightly	Moderately	A lot
Informative	0	9,8	**74,5**	15,7
Educational programs	12,7	**46,1**	**33,3**	7,8
Assessment/Diagnosis	**43,1**	**46,1**	9,8	1
Corrective/Conflict resolution	0	16,7	**54,9**	**28,4**

Bold values represents the explanation of results.

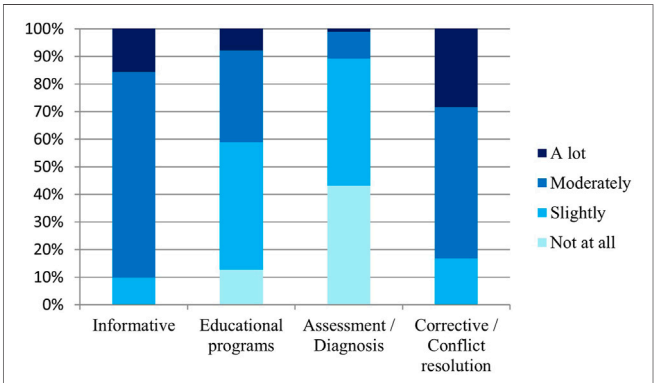


FIGURE 1 | Distribution of intervention models aimed at students.

TABLE 3 | Frequencies obtained for the different types of intervention with families.

	Not at all	Slightly	Moderately	A lot
Informative	2,9	**36,3**	**52**	8,8
Educational programs	**50,0**	**43,1**	5,9	1
Information request	12,7	**38,2**	**41,2**	7,8

Bold values represents the explanation of results.

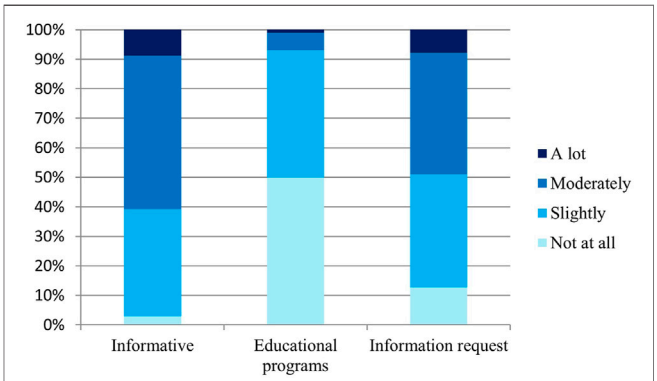


FIGURE 2 | Distribution of intervention models aimed at families.

skills (33.33%), emotional education (31.37%) and affective-sexual education (30.39%). In the fourth year of secondary school, the highest priority was guidance in academic-professional pathways and actions aimed at facilitating decision-making, reaching 87.25% combined. Most participants answered by selecting both responses. Coexistence was chosen as a priority by 5.88%.

The tutoring activities aimed at students were classified into four models or typologies: 1) Informative: to provide and exchange information. 2) Educational programs: to develop training through programming with objectives, content, activities and evaluation. They have an educational and, therefore, preventive purpose. 3) Assessment/Diagnosis: to use assessment instruments to detect needs and collect data in order to direct the intervention, refer to other guidance services, etc. 4) Corrective/Conflict resolution: to act in a corrective manner at specific times after the appearance of a coexistence conflict or poor academic performance.

The results showed the distribution of the procedures used, highlighting those aimed at reporting and correcting problematic situations. The application of educational programs remained at intermediate frequency levels, and very few evaluative or diagnostic actions were carried out by tutors (**Table 2** and **Figure 1**).

Regarding families, results were obtained pertaining to three types of procedures: activities with the purpose of informing, educational actions through programs and information requests. High levels were not reached for any of the three typologies, indicating a rather moderate consideration of interventions aimed at the families of students. However, medium frequencies were found for the two types that refer to information exchange, preferably developed through group meetings and individual interviews. There were very few educational program actions (**Table 3** and **Figure 2**).

Examining the individual interviews with families, as it is the most common way of relating to teachers, there were rather wide gaps with respect to the topics discussed. Ninety-six percent (96%) of the counselors stated that teacher-tutors addressed academic topics in interviews, highlighting the information on grades. A total of 89.2% also selected the topic of conflict management. A total of 79.4% addressed training pathways and career opportunities in interviews. A total of 65.7% dealt with aspects of socioaffective development, well-being, social integration and motivation. It is estimated that less than half of the guardians (41.2%) addressed household habits and possible family needs. A total of 21.6% reported that tutors acknowledge, in interviews, students' best and most positive competencies. Finally, only 8.8% answered that they talk about values and personality development.

Regarding the strategies and instruments used to obtain information about the aptitudes, interests and competencies of students, to provide professional advice, the results provide very interesting information. There was not a single procedure that reached 70% affirmative answers; therefore, there is great heterogeneity among procedures that tutors use to get to know their students in vocational guidance. The three data sources most considered by the tutors were qualifications (68.6%), informal learning (64.7%) and systematic observation (60.8%). Forty-eight percent stated that personal interviews were conducted with students. A total of 35.3% felt that tutors conducted interviews with families for the purposes indicated. A total of 34.3% included tests as sources of information. Finally,

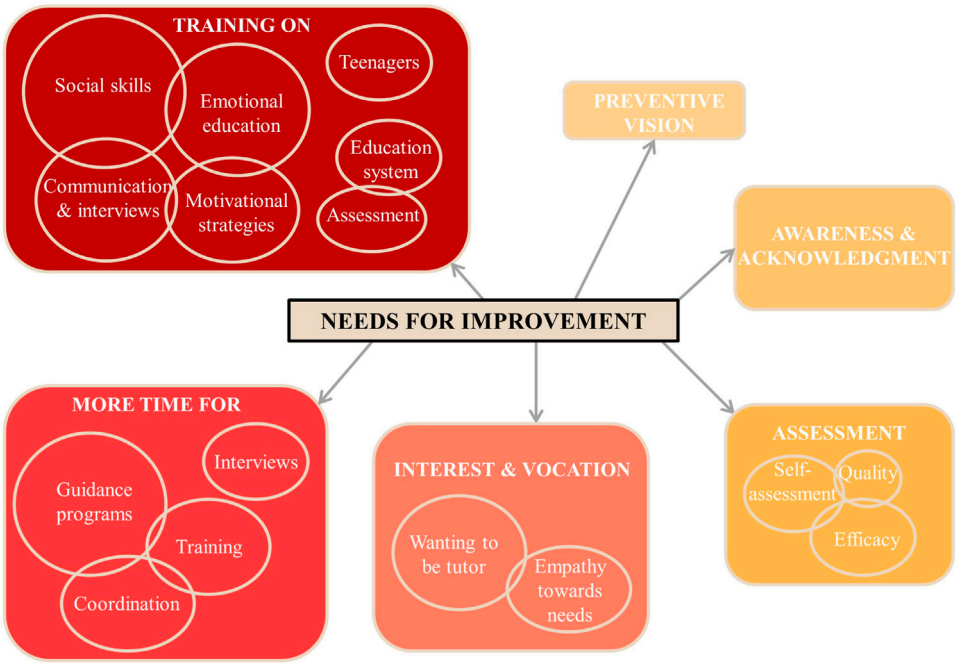


FIGURE 3 | Needs for improvement in the tutorial actions developed by teachers.

the narratives developed by the students, such as portfolios and blogs, were only used by 16.7% of the teachers.

The presentation of the results is concluded with a concise and meaningful assessment of tutorial action in secondary education in relation to effectiveness and the needs detected. The counselors felt that the involvement and commitment of the tutors, in the performance of the functions that are the object of this study, were very diverse. Added to the dispersed responses is that 18.6% gave the tutors at their educational centers a score less than 5 on a scale from 0 to 10, while 26.5% rated them between 5 and 6. Forty-five percent chose 7 or 8 as an answer, 9.8% provided a score of 9 for tutor involvement, and no participant selected the maximum score of 10. Furthermore, the assessment of the regulations on counseling and tutoring in effect in Spain was poorly rated for all items included in the questionnaire. A total of 64.7% of the participating counselors believed that the regulations are not sufficient, 65.7% stated that they lack clarity, and 73.5% rated them as unhelpful for practice.

The last phase of the analysis refers to the data collected through open-ended questions on the questionnaire, related to the needs that the counselors detect and their proposals for improvement (RQ3). There were six categories of perceived interconnected needs, as seen in **Figure 3**. In order of priority from highest to lowest, they were training, time, interest, evaluation, recognition and prevention. Some of them, in turn, included subcategories, presented with their respective saturations and interconnections.

The findings highlight the need for training, above all others. The counselors believed that teachers should be better trained in pedagogy and psychology content related to coexistence skills, emotional education, communication strategies, interviews, the education system, adolescence, motivation and evaluation: "They greatly need continuing education" (P97), "They lack communication skills, empathy, emotional management, . . . " (P11), "It is more difficult for them to conduct 1 hour of tutoring than the 19 h of class" (P84), and "Very few educational professionals seem to be truly prepared to work with adolescents and understand their personal, social, emotional, issues etc." (P17). The need with the second-greatest presence was time dedicated to counseling activities by teachers. They feel that more time should be allocated within the school day to develop tutoring programs, conduct interviews, coordinate and train: "More hours set aside for tutoring" (P15), "They need more time, not more resources" (P46), and "The teachers have no time to coordinate among themselves" (P31). The third was awareness of the educational function of teachers, which is evident in career interest and development. A significant number of counselors felt that some of the teaching staff did not want to be tutors and performed this function without motivation: "Some tutors aren't motivated" (P68), "Predisposition of teachers" (P90), and "Tutors lack the ability to empathize and connect with the students" (P54).

The following thematic categories differed significantly from the first in saturation. Fourth, there was the need to increasingly evaluate tutorial action, which lacks formal means and instruments and is relegated to an assessment of satisfaction and opinions in many cases: "There is little evaluation, only follow-up in the meetings with tutors" (P6), "It is not formally evaluated" (P27), and "With opinion polls to students" (P20). The fifth need refers to the recognition of the role of the guardians. This category is defined as poor clarification and appreciation of their pedagogical work by their own selves, by the students and families: "Lack of specific recognition of the tutorial function" (P12), "A tutor's work is not valued, neither economically nor in

schedule or in scale" (P101), "I wish to add that the figure of the tutor is voluntary and recognized" (P68). Finally, a sixth category appears with very slight saturation but is included in the results due to its theoretical significance. It is the need to appreciate the positive aspects in students and to intervene preventively. Some participants mention that the vision that many tutors have regarding guidance is exclusively academic and reduced to solving conflicts: "Having more of an interest in aspects other than academics" (P88) and "Seeing and highlighting your strengths and using them to your advantage and enhancing them in the classroom" (P7).

All the results presented correspond to the research questions and the objectives of the research. However, as part of the more open narratives provided by the counselors, ideas were recorded that had not been considered in the design of this study but that have contributed new research interest. The statements by the counselors invited a reflection on their own role as teacher-tutors and whether it is appropriate to invest in their psychopedagogical training for that role or to create a specific and more competent support network. A third approach was also proposed, i.e., to provide each group with more than one tutor and then there would be no distinction between tutor and nontutor teachers. The textual excerpts included "There are magnificent tutors who do their work very well, and others who do not. However, the system does not facilitate the latter to improve. In our center, most programs are carried out by outside parties, who do their job well" (P46), "It is essential to study this and definitely question the actual usefulness of this weekly space (in reference to the hour of tutoring)" (P70), "The fact that not every teacher is a tutor leads to a uneasiness every year that is pretty big depending on the serious headaches caused by the group. It would be interesting if all teachers, with the exception of department heads who perform other administrative functions, were tutors, and if there were 2 or 3 tutors in the most complex groups. Tutoring is so fundamental that we have to take better care of so that it provides a real approach and guidance for the students and their families" (P84).

DISCUSSION AND CONCLUSION

This study contributes a different view to the current research on tutorial action by contributing data obtained directly from those responsible for counseling. The aim is to complement the conclusions provided by the studies conducted from the perspectives of teaching faculty (López, 2013; Vélaz-de-Medrano et al., 2018), students (Amor & Serrano, 2020) and families (Castro et al., 2015). However, considering their exploratory aims, these conclusions are the first step towards a more detailed line of research.

In the first phase of the analysis, significant emerging relationships were found between certain characteristics of the participating counselors and some of the variables studied. Previous hypotheses had not been considered; therefore, the findings invite reflection on the professional meaning of these relationships and the definition of hypotheses for subsequent confirmatory studies. Such heterogeneity and apparent randomness are perceived in the relationships found, which can be linked to lack of time, excessive tasks and disproportionate student-teacher ratios. Additionally, there is a lack of clarity in guidance functions and regulations regarding the prioritization of the content and diversity of intervention models. Some issues also arise based on the interpretation of the different training profiles leading to the professional role of counselor. The most significant difference regarding the introduction of guidance content in teaching has been found in relation to gender and is in favor of women. This result agrees with the conclusions of OECD (2009) when concluding that women focus their teaching centered on students with greater consideration to their personal and cooperative development, and less to the transmission of knowledge. That is, in essence, counselors are professionals with vastly different personal and training backgrounds who develop the same professional path, coinciding with other studies carried out with educational professionals (Sutton and Fall 1995; Oyifioda and Iornenge, 2020). Is this heterogeneity positive in the professional role of educational guidance?

The first conclusion refers to the beliefs that counselors express in this study regarding the educational, not didactic, work of teachers in their schools. In general, guidance actions developed thus far are insufficient, unmotivated in many cases and very heterogeneous. There is a lack of consensus regarding the organization of content, use of strategies, selection of evaluation instruments and implementation of actions. In addition, counselors perceive wide differences between the actions of tutors and other teachers, although all teachers should develop guidance work, according to the rules and theory (González-Benito and Vélaz de Medrano, 2014).

The problems of demotivated tutors and the lack of recognition coincide with the needs detected in the research that gather their opinion in this regard (López, 2013). On the other hand, some of the needs found in this study correspond to the mismatch between the real and ideal dedication of tutors to guidance work found by Vélaz-de-Medrano et al. (2018). In their conclusions, several causes are mentioned that are related, in a circular way, with the dissatisfaction expressed by tutors: high level of self-demand, insufficient training, lack of time and complex contexts. In the results of this research, there are also several categories of needs interconnected by counselors. As seen in **Figure 3**, there is full agreement regarding the gaps in training and time. However, counselors do not perceive the high level of self-demand to which teacher-tutors feel subjected in conjunction with little training. The possibility emerges that there is not enough mutual understanding and that, in a complex and demanding context, both types of professionals need more tools for communication, empathy and coordination. In addition, from the perspective of counselors, other deficiencies have been detected that, although less saturated, are also relevant. They refer to the fact that the evaluation of interventions is not very useful; it is not aimed at assessing quality but satisfaction and is infrequent. Additionally, teachers do not conceive of tutoring from a preventive perspective but rather under the concept of interventions aimed at "putting out fires", solving problems and informing families about academic issues.

The need to increase pedagogical and guidance training among all secondary school teachers in Spain fully coincides with the

guidelines indicated by Imbernón (2019), among others, for a long time. It is an argument for which the theoretical visions of two areas of pedagogical knowledge are fully integrated: guidance and didactics. However, this has not yet been reflected in the plans for the initial training of secondary school teachers. These plans continue to be more concerned with the expansion of curricular content and specialization pathways, as well as digital innovation, than with revaluing the humanization of the educational process (Loughran and Hamilton, 2016).

When evaluating the coincidences and discrepancies between perspectives from which tutorial action is studied, some doubts also emerge that invite further research and contrasts. For example, what topics are covered in the interviews between tutors and relatives of students? None of the counselors left this question blank, obtaining yes/no answers for all the proposed items in all cases. That is, the counselors show their beliefs about something that happens without being present. The results give much greater presence to some topics in the interviews, such as conflict and academic pathways, to the detriment of other topics, such as positive competencies and personality development. However, by inference, a finding is also obtained related to the authority shown by the counselors in regard to commenting on the most purely pedagogical actions performed by teachers, using information from their perception and interpretation. This idea is consistent with the conclusions reached by Hernández and Mederos (2018) in their study on the self-definition of expert counselors in educational institutions.

When interpreting the results by nuclei of interest, several specific conclusions are also interesting. One of the concerns that motivated the design of this research is obtaining knowledge regarding what guidance content is taught by tutors and whether all teachers integrate such content in the subjects they teach. The purpose of accounting for what time is spent, with what order of priority and in which courses has been achieved, considering the results with the prudence that corresponds to the size of the sample. The first paradox generated is that the content was most effective for students from first to third grade, that is, training in social skills, improving coexistence and emotional education, areas in which counselors refer the most need for teacher training. The second arises when verifying that some topics, such as education in values, affectivity and sexuality, addiction prevention, communication or entrepreneurship, are not very visible. Specifically, a significant number of participants alluded to the lack of communication skills exhibited by tutors in particular and by teachers in general but hardly considered communicative competence in the programming of tutorial action aimed at students. Communication seems to be eternally forgotten in training plans, both in secondary education and in teacher training itself (Chan et al., 2011).

The prioritization of content related to vocational guidance in the fourth year is not a surprise and is logical. However, the fact of being practically absent in previous courses and concentrating on making academic and professional decisions leads us to think that students do not learn self-determination progressively; they do not become competent in decision-making. Based on the results obtained, this topic is still far from being integrated by schools into career guidance content (Andrei and Solberg, 2019; Babarović et al., 2020).

Relatedly, there is little intervention through educational programs. This reality has been confirmed in two ways: through questions about the most commonly used intervention models with students and families and through open-ended questions about improvement needs. What is worrying is that guidance actions lack preventive and developmental vision. The counseling model prevails, with great dedication to managing conflicts and needs as they arise. This detected reality is far from the authentic curricular integration of counseling (Sultana R, 2018; Boland et al., 2019).

In relation to the assessment instruments, the least used instruments are those with which students assume the leading role and lead them to reflect, such as the construction of narratives. The systematic evaluation to better know students and accompany them in decision-making is not very common, based on intermediate frequencies, although the use of observation is quite widespread. The two main conclusions for this section are the heterogeneity of processes and the majority use of information from daily and informal treatment. Both lead to inferences of spontaneity and a lack of evaluative sense. They correspond to the detected need for teacher training in evaluation skills. In this aspect, it would be desirable to improve the assessment processes developed, both to know the student in transversal skills and socioemotional development and to conduct formative assessments in subjects.

The use of interviews was studied specifically because of its transcendence within the guidance processes developed by teaching faculty. Interviews with students are not as common as could be expected; their use is estimated below 50% as a vocational guidance tool, and the percentage decreases in relation to their use with families. In addition, counselors believe that the topics addressed are grades, academic difficulties, pathways and conflict management, to a greater extent. If one considers the use of interviews with approximately half of the high school students and that the topics on emotional, social development, transversal competences and values occupy between 8 and 60%, it can be concluded that approximately three-quarters of the students never have a personal and private conversation with their tutor about the most valuable counseling content. This is an inference that requires more scientific evidence but that, within the framework of this study, invites reflection on what may be the greatest lack of the system in relation to what adolescents most need (Aguiar and Conceição, 2015; Park et al., 2018).

Finally, within the needs and shortcomings described by the counselors, the regulations from the Spanish educational administration stands out as being not very useful for practice, insufficient and lacking clarity. This is a common description in counselor meetings and professional journals (APOEX, 2019). With this study, quantitative data were obtained that support this opinion distributed throughout the national territory and with a sample of highly experienced counselors.

As it is an exploratory study of a descriptive nature, there are several limitations. First, the need for interpretation of qualitative data. Second, the size of the sample that makes it impossible to generalize the results. To continue advancing in the line and complement this study, the prospective research line focuses on conducting in-depth interviews with counselors, tutors and students. It is intended to combine a dual approach, at the national and regional levels, linking the

results of the exploratory study already conducted with interviews. Another objective is to enrich collaborative continuity with professional entities of psychopedagogy and initiate relationships with public bodies competent in education. The aim is to obtain support for the research line and to transfer the conclusions to future projects related to regulations for tutorial actions.

New questions also arise: How can the theoretical vision of tutoring and practical action in the current system be approached with maximum sustainability? Currently, in the spanish education system, counselors have the function of coordinating and overseeing tutorial actions and are overwhelmed by the lack of regulatory concretion and the ratios with which they work. Can direct intervention by counselors be increased in a system in which teachers lack sufficient training and dedication to tutoring, or should guidance training for teachers be prioritized and their dedication to teaching be increased for these purposes? Ultimately, the possibility emerges that training in emotional and social competencies, drug addiction prevention or affective-sexual education, among other content, can be addressed by outside agents who are experts in these matters or that a new professional profile is generated within the educational institutions in charge of these functions that reinforce guidance departments.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Aragón Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHORS' CONTRIBUTIONS

Conceptualization, CF-L, AM, and AP; formal analysis, CF-L, AM, AP, and AR; methodology, AR and CF-L; writing—review and editing, CF-L, AM, AP, and AR.

ACKNOWLEDGMENTS

We express our appreciation to the participating entities (AAPS and COPOE) and the Office for the Transfer of Research Results (OTRI in Spanish) of the University of Saragossa, for facilitating the collaboration agreement. We also thank the University of Saragossa, in particular the Vice-Chancellor of Technology Transfer and Innovation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2021.737163/full#supplementary-material

REFERENCES

Aguiar, F. H. R., and Conceição, M. I. G. (2015). Career Guidance and Health Promotion with Adolescents: Keys for Intervention Programs. *Int. J. Educ. Vocat Guidance* 15 (3), 221–236. doi:10.1007/s10775-015-9311-9

Aguilar Gavira, S., and Barroso Osuna, J. (2015). La triangulación de datos como estrategia en investigación educativa. *Pixel-Bit* 47, 73–88. doi:10.12795/pixelbit.2015.i47.05

Amor Almedina, M. I., and Serrano Rodríguez, R. (2020). Las competencias profesionales del orientador escolar: el rol que representa desde la visión del alumnado. *rie* 38 (1), 71–88. doi:10.6018/rie.321041

Andrei, A., and Solberg, S. V. (2019). "Personalized Career and Academic Planning-A Promising Policy and Practice," in Annual Conference of the International Association of Vocational and Educational Guidance Bratislava, Slovakia. 2019-09-09 - 2019-09-12, Available at: https://hdl.handle.net/2144/40489.

APOEX (2019). *Conclusiones III Congreso Internacional de Orientación y X Encuentro Estatal de Orientación Educativa: "Orientación educativa: Nuevas realidades, apasionantes retos"*. Badajoz, Spain: Asociación Profesional de Orientadores de Extremadura.

Auger, R. W., Abel, N. R., and Oliver, B. M. (2018). Spotlighting Stigma and Barriers: Examining Secondary Students' Attitudes toward School Counseling Services. *Prof. Sch. Couns.* 22 (1), 2156759X1881127. doi:10.1177/2156759x18811275

Babarović, T., Dević, I., and Blažev, M. (2020). The Effects of Middle-School Career Intervention on Students' Career Readiness. *Int. J. Educ. Vocat Guidance* 20 (2), 429–450. doi:10.1007/s10775-019-09411-5

Benigno, S. (2017). Counselor Perceptions: Let Us Do Our Job!. *Jel* 6 (4), 175–180. doi:10.5539/jel.v6n4p175

Boland, L., Légaré, F., Carley, M., Graham, I. D., O'Connor, A. M., Lawson, M. L., et al. (2019). Evaluation of a Shared Decision Making Educational Program: The Ottawa Decision Support Tutorial. *Patient Educ. Couns.* 102 (2), 324–331. doi:10.1016/j.pec.2018.09.008

Cano-Escoriaza, J., Hernández, S. O., and Pascual, A. C. (2019). Work Values of Spanish University Students: a Follow-Up Study. *Vocations Learn.* 12, 425–439. doi:10.1007/s12186-019-09221-4

Carey, J. C., Martin, I., Harrington, K., and Trevisan, M. S. (2018). Competence in Program Evaluation and Research Assessed by State School Counselor Licensure Examinations. *Prof. Sch. Couns.* 22 (1), 1–11. doi:10.1177/2156759x18793839

Castellano, M. E., Richardson, G. B., Sundell, K., and Stone, J. R., III (2017). Preparing Students for College and Career in the United States: the Effects of Career-Themed Programs of Study on High School Performance. *Vocations Learn.* 10, 47–70. doi:10.1007/s12186-016-9162-7

Castro, M., Expósito, E., Lizasoain, L., López, E., and Navarro, E. (2015). Acciones y actitudes diferenciales de los tutores y su relación con la participación de las familias. *Participación Educativa, Revista Del. Consejo Escolar Del. Estado* 4 (7), 29–37. doi:10.4438/1886-5o97

Chan, R., Lau, P., and Yuen, M. (2011). Interrelationships Among Teacher Care, Students' Life Skills Development, and Academic Achievement: Implications for School Guidance Work. *Asian J. Counselling* 18 (1), 63–94. Available at: https://www.hkier.cuhk.edu.hk/journal/document/AJC/ajc_v18n1-2_63-94.pdf.

Costa Jiménez, P., Martínez Ramón, J. P., Ruiz Esteban, C., and Méndez Mateo, I. (2020). Análisis de la orientación académica y profesional en Educación Secundaria y su adaptación al alumnado con necesidad específica de apoyo educativo. *Reifop* 23 (1), 175–191. doi:10.6018/reifop.373701

Council of the European Union (2021). Council Resolution on a Strategic Framework for European Cooperation in Education and Training towards the European Education Area and beyond (2021-2030), (2021/C 66/01). *Official J. Eur. Union*. doi:10.1787/ba121fd9-en

ETF (European Training Foundation). (2020). International Trends and Innovation in Career Guidance. Vol. I. Available at: https://www.etf.europa.

eu/en/publications-and-resources/publications/international-trends-and-innovation-career-guidance-0

European Commission (2020). "Lifelong Guidance Policy and Practice in the EU: Trends, Challenges and Opportunities,". Final report. ISBN 978-92-76-17340-3. doi:10.2767/91185

Flick, U. (2009). *An Introduction to Qualitative Research*. London, UK: SAGE Pub.

González-Benito, A., and Vélaz de Medrano, C. (2014). *La acción tutorial en el sistema escolar*. Madrid , Spain: UNED.

Haug, E. H., and Plant, P. (2016). Research-based Knowledge: Researchers' Contribution to Evidence-Based Practice and Policy Making in Career Guidance. *Int. J. Educ. Vocat Guidance* 16 (1), 137–152. doi:10.1007/s10775-015-9294-6

Hearne, L., and Neary, S. (2021). Let's Talk about Career Guidance in Secondary Schools! A Consideration of the Professional Capital of School Staff in Ireland and England. *Int. J. Educ. Vocat Guidance* 21 (1), 1–14. doi:10.1007/s10775-020-09424-5

Hernández Rivero, V., and Mederos Santana, Y. (2018). Papel del orientador/a educativo como asesor/a: funciones y estrategias de apoyo. *REOP* 29 (1), 40–57. doi:10.5944/reop.vol.29.num.1.2018.23293

Imbernón Muñoz, F. (2019). La formación del profesorado de secundaria: la eterna pesadilla. *Revista de Curriculum y Formación Del. Profesorado* 23 (3), 151–163. doi:10.30827/profesorado.v23i3.9302

Jariot, M. (2010). Cómo implicar al equipo docente en el proceso de orientación. Un ejemplo de superación del enfoque de servicios, *REOP* 21 (1), 45–58. doi:10.5944/reop.vol.21.num.1.2010.11508

Johnson, J., Rochkind, J., and Ott, A. (2010). Why Guidance Counseling Needs to Change. *Educ. Leadersh.* 67 (7), 74–79.

López, E. (2013). Aproximación a la percepción y satisfacción del profesor tutor de Secundaria Obligatoria respecto a su labor. *Revista de Investigación en Educación* 11 (1), 77–96. Available at: https://reined.webs.uvigo.es/index.php/reined/article/view/163.

Lorenzo Vicente, J. A., Muñoz Galiano, I. M., and Beas Miranda, M. (2015). Modelos de formación inicial del profesorado de Educación Secundaria en España desde una perspectiva Europea. *Rev. Complut. Educ.* 26 (3), 741–757. doi:10.5209/rev_rced.2015.v26.n3.44866

Loughran, J., and Hamilton, M. L. (2016). *International Handbook of Teacher Education (vol 1)*. Springer, Singapore.

Louws, M. L., Meirink, J. A., van Veen, K., and van Driel, J. H. (2017). Exploring the Relation between Teachers' Perceptions of Workplace Conditions and Their Professional Learning Goals. *Prof. Dev. Edu.* 43 (5), 770–788. doi:10.1080/19415257.2016.1251486

Neumann, K., Kind, V., and Harms, U. (2019). Probing the Amalgam: the Relationship between Science Teachers' Content, Pedagogical and Pedagogical Content Knowledge. *Int. J. Sci. Edu.* 41 (7), 847–861. doi:10.1080/09500693.2018.1497217

OECD (2019). *Creating Effective Teaching and Learning Environments: First Results from TALIS 2009*. Paris, France: OECD Publishing, Available at: https://www.oecd.org/education/school/43023606.pdf.

Oyifioda, B., and Iornenge, I. (2020). The Influence of Demographic Variabes on Teachers´ Occupational Interests in Public Secondary Schools in North Central Nigeria. *Int. J. Edu. Eval.* 6 (5), 14–23. Available at: https://iiardpub.org/journal/?j=IJEE.

Park, J.-H., Rojewski, J. W., and Lee, I. H. (2018). Determinants of Adolescents' Career Development Competencies in Junior Secondary Schools of South Korea. *Int. J. Educ. Vocat Guidance* 18 (1), 1–25. doi:10.1007/s10775-017-9342-5

Perry, J. C., and Shannon, L. (2017). How Vocational Psychologists Can Make a Difference in K-12 Education. *Int. J. Educ. Vocat Guidance* 17 (1), 97–115. doi:10.1007/s10775-016-9335-9

Ries, F., Cabrera, C. Y., and Carriedo, R. G. (2016). A Study of Teacher Training in the United States and Europe. *Ejsbs* 16 (2), 86–110. doi:10.15405/ejsbs.184

Rodríguez Álvarez, P., Ocampo Gómez, C. I., and Sarmiento Campos, J. A. (2018). Valoración de la orientación profesional en la enseñanza secundaria postobligatoria. *rie* 36 (1), 75–91. doi:10.6018/rie.36.1.285881

Rodríguez Martínez, A., Sierra Sánchez, V., Falcón Linares, C., and Latorre Cosculluela, C. (2021). Key Soft Skills in the Orientation Process and Level of Employability. *Sustainability* 13 (6), 3554. doi:10.3390/su13063554

Silva, I., Diniz, J., and Zeichner, K. (2017). *Justicia social. La dimensión olvidada de la formación docente*. Santiago: Mutante Editores.

Strauss, A., and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedure and Techniques*. NewburyPark, London, UK: Sage.

Strauss, A. (1987). *Qualitative Analysis for Social Scientists*. New York, USA: Cambridge University Press.

Sultana, R. G. (2014). Pessimism of the Intellect, Optimism of the Will? Troubling the Relationship between Career Guidance and Social justice. *Int. J. Educ. Vocat Guidance* 14 (1), 5–19. doi:10.1007/s10775-013-9262-y

Sultana, R (2018). *Enhancing the Quality of Career Guidance in Secondary Schools. . . a Handbook. MyFuture Erasmus+ Project Output*. Msida, Malta: University of Malta.

Sutton, J. M., and Fall, M. (1995). The Relationship of School Climate Factors to Counselor Self-Efficacy. *J. Couns. Dev.* 73 (3), 331–336. doi:10.1002/j.1556-6676.1995.tb01759.x

Vaishnavi, J., and Kumar, A. (2018). Parental Involvement in School Counseling Services: Challenges and Experience of Counselor. *Psychol. Stud.* 63 (4), 359–364. doi:10.1007/s12646-018-0463-9

Vélaz-de-Medrano, C., González-Benito, A., and López-Martín, E. (2018). Assessment of the Level of Performance of Tutoring in Secondary Education: Tutors´ Own Perceptions. *Revista de Educación* 382, 107–132. doi:10.4438/1988-592X-RE-2018-382-394

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

CONCEPTUAL ANALYSIS
published: 20 October 2021
doi: 10.3389/fpsyg.2021.725407

A Conceptual Framework for Understanding Variability in Student Perceptions

Bilge Gencoglu, Michelle Helms-Lorenz, Ridwan Maulana and Ellen P. W. A. Jansen*

Department of Teacher Education, University of Groningen, Groningen, Netherlands

Student perceptions using surveys are frequently used to measure student perceptions of teachers' teaching quality in secondary and higher education. Research shows that the variance in student perceptions exists within a class and between countries. However, the influence of individual- and cultural-level factors on the variance of student perceptions is less well studied. More insights are needed to understand the mechanisms underlying the variance in student perceptions in-depth. Insights into determinants of student perceptions of teaching quality could become valuable toward understanding school-related outcomes. A conceptual framework is put forward in this study to enhance our understanding of manifestations of student perceptions of teaching quality. It is suggested that value orientations at the individual- and cultural-level as well as social desirability may play a role in understanding student perceptions of teaching quality. Understanding students' individual and collective perceptions of teaching quality can contribute to teachers' sense-making of their student evaluations. It is argued that this understanding could contribute to enhancing the development of teaching quality and ultimately education quality.

Keywords: student perception, teaching quality, value orientation, value congruence, variety of perception, social desirability

OPEN ACCESS

Edited by:
Kathryn Holmes,
Western Sydney University,
Australia

Reviewed by:
Erik A. Ruzek,
NWEA, United States
Patcharin Panjaburee,
Mahidol University, Thailand

****Correspondence:***
Bilge Gencoglu
b.gencoglu@rug.nl

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *15 June 2021*
Accepted: *27 September 2021*
Published: *20 October 2021*

Citation:
Gencoglu B, Helms-Lorenz M, Maulana R and Jansen EPWA (2021) A Conceptual Framework for Understanding Variability in Student Perceptions. Front. Psychol. 12:725407. doi: 10.3389/fpsyg.2021.725407

INTRODUCTION

Student perceptions of teachers' teaching quality are important for students' school-related outcomes, such as academic motivation and engagement (Maulana et al., 2016; Maulana and Helms-Lorenz, 2016; Inda-Caro et al., 2019). Although the importance of student perceptions for school outcomes is largely acknowledged by scholars, more insights are needed to understand the factors that influence student perception besides the teacher's behavior in the classroom. Student perceptions of the same teacher's teaching quality vary within a class (between-student variations; Maulana et al., 2015). The variety in student perceptions of the same teacher suggests that there are individual factors affecting student perceptions. Similarly, student perceptions vary across countries (André et al., 2020). The variety in student perceptions within and across countries may be an indication that there are cultural factors affecting student perceptions. This conceptual paper explores value orientations as an underlying individual- and cultural-level factor explaining variations in student perceptions.

Explorations of the variations in student perception have the potential to enrich our understanding of the nature of the information contained in student ratings. However, variance among students within the same classroom is frequently treated as an error or nuisance variation (e.g., Meade and Eby, 2007). In other words, it is assumed that "each student would assign the same rating, such that the responses of students in the same class would be interchangeable" (Lüdtke et al., 2006, p.217). Compared to the immense body of research focusing on the mean levels of aggregated student ratings at the class-level, much less attention has been paid to the degree to which students within a class agree in their perceptions (within-class variety, e.g., Lüdtke et al., 2006; Schweig, 2016). These studies investigate whether within-class variability[1] has predictive power for learning and achievement outcomes (Schweig, 2016; Schenke et al., 2017, 2018; Bardach et al., 2019b, 2021).

The variety of perceptions can be delineated by the psychological processes described by scholars of person perceptions. The rational-emotive behavior (REB) model (Ellis, 1962) describes how the interpretation of the same observed behavior by two different people can lead to different perceptions. One of the primary tenets of the REB model is that thoughts, feelings, and behaviors interact and affect each other (Ellis, 1962). Ellis (1991) asserts that thinking affects and creates feelings and behaviors; at the same time, emotions influence thoughts and behaviors, and behaviors have an impact on thoughts and feelings. Consequently, if one of these processes is altered, the other processes are influenced dynamically as well (Banks and Zionts, 2009). Because individual experiences and reactions to experiences are different, the perceptions will differ, too. The hypothesis of equal perceptions of teaching quality is highly unlikely given the psychological mechanisms put forward by Ellis's model.

Ellis (1974) posits an ABC theory of disturbance. A is defined as an Activating event or activating experience; B is defined as a Belief system; and C is defined as emotional and behavioral Consequences. These consequences may either be (1) healthy as a result of rational perceptions or (2) self-defeating as a result of irrational or distorted perceptions (Banks and Zionts, 2009). Based on REB, Clawson (2012) proposes that observed behavior is compared to a personal set of values, assumptions, beliefs, and expectations about the way the world is or should be. This comparison leads one to make conclusions and interpretations about the observed behavior, which determines the perception of what is observed. The role of beliefs is hypothesized as mediating the relationship between external events and emotional consequences. Irrational beliefs lead to disturbing emotions, such as depression, fear, anger, and negative behavioral reactions, such as withdrawal and impulsivity (Bernard, 1990). Rational beliefs, on the other hand, generally result in moderate emotions that foster goal attainment and life satisfaction (Smith, 1982). In other words, an individual's emotional and behavioral response is thought of as caused by an external event, yet it is the result of a combination of an external event and the processing of the information by the individual's belief system (Ellis, 1962).

Applying the REB model to student evaluations of teaching quality implies that evaluations do not only consider the observed external event, that is, teaching quality, but also incorporate the processing of the information by students, that is, personal set of belief systems. A value-belief-norm theory suggests that individual values lead to beliefs, which in turn help to form personal norms (Stern et al., 1999; Stern, 2000). Given that values operate to shape beliefs and norms, they are also involved in student evaluations of teaching quality. In this conceptual paper, values are explored to understand student's individual and collective perceptions of teaching quality. Student evaluation of teaching quality is treated as an instance of perceptions and value orientations as underlying reasoning behind perceptions.

Teaching quality can be conceptualized in many different ways (Helms-Lorenz and Visscher, 2021). Scheerens (2016) postulated a distinction between the pro-active (the preparation activities and prerequisites involved before a lesson is executed), interactive [the (in)visible interactions during the lesson], and retro-active (the evaluation of the conducted lesson and of student learning after the execution of the lesson) aspects of teaching. Although manifest teaching behavior in the classroom is viewed as a proxy of teaching quality reflecting the three aspects of teaching (Helms-Lorenz and Visscher, 2021), student perceptions of teaching quality address the classroom teaching behavior during the lesson because student perceptions are linked with teachers' act in the classroom. Various frameworks have been developed to study classroom teaching behavior and its relation to students' school-related outcomes (Creemers, 1994; Scheerens et al., 2007; van de Grift, 2007; Creemers and Kyriakides, 2008; Pianta et al., 2008; Hattie, 2009; Pianta and Hamre, 2009; Sammons and Bakkum, 2011; Danielson, 2013). Although these frameworks lead to several measurement instruments varying in terms of structure and their main underlying models and conceptualizations, at least six teaching behavior domains were revealed showing a relationship with students' learning and outcomes: Providing a Safe and Stimulating Learning Climate, Efficient Classroom Management, Clarity of Instruction, Activating Learning, Adaptive Teaching, and Teaching Student Learning Strategies. Also, teacher-student interpersonal relationship has been shown to be an important determinant of the learning processes of students (van Tartwijk et al., 1998; den Brok, 2001). For example, the framework developed by Hamre et al. (2013), which focuses on teacher-student interactions, proposes a three-domain structure, namely, emotional supports, classroom organization, and instructional supports. Although each framework has significant contributions in the literature, choosing a specific framework of teaching quality might narrow down the scope of our proposed framework. It is thus noteworthy to mention that the conceptualization of teaching quality is not restricted to a specific theoretical framework.

In the following sections, values are defined and individual value orientations are reviewed. The concept of value (in) congruence is presented with regard to interactions resulting

[1]The terms "within-class consensus" and "heterogeneity in student perceptions" were used to specify the degree of within-class variability.

from similar and dissimilar values and applied to student perceptions of teaching quality. Then, cultural value orientations are put forward and applied to student perceptions of teaching quality. It is followed by unpacking the concept of social desirability, which addresses the tendency to share similar ideas with others. By proposing a conceptual framework, we conclude that value orientations, value congruence, and social desirability provide rich background information that might allow predictions and understanding of the student perceptions variations within and between cultures.

VALUES

Values are defined as deep, enduring, and guiding principles that influence people's behavior (Rokeach, 1973; Schwartz, 1992; Parashar et al., 2004). They are desirable, abstract goals that guide beliefs, attitudes, perceptions, and behaviors (Rokeach, 1973; Schwartz, 1992) that reflect the way one wants the world to be. While norms, attitudes, and specific goals depend on specific situations, actions, or objects, values transcend specific situations (Schwartz, 1992). Unlike traits, orientations, and interests, values serve as criteria or standards to provide social justification for choices and behaviors (Roccas et al., 2002; Sagiv, 2002). Values help to shape personal and collective preferences about what is important. In other words, values guide individuals and societies about how to act in particular situations.

Schwartz (1992, 1994) developed two different value theories, one regarding individual differences in values and the other regarding cultural differences. The first theory assumes that values are ordered by subjective importance and form a unique system of value orientations at the individual level (Rokeach, 1973; Schwartz, 1992). The second theory assumes that values are ordered by hierarchical social systems, reflecting the value orientations at the cultural-level (Schwartz, 2006). The two proposed orientations reflect that the culture-level processes are different from individual-level processes (Schwartz, 1999; Fischer et al., 2010). By doing so, Schwartz followed the same tradition established by Hofstede (1980) who argued that individual and cultural levels are not isomorphic (Fischer et al., 2010). Schwartz proposed different structures at the two levels arguing that the characteristics that discriminate among societies and those that discriminate among individuals are unlikely to be the same (Schwartz, 1994). In the following sections, we explore the possible influence of individual and cultural values on student perceptions.

Individual Values

At the individual level, Schwartz's (1992) postulates ten basic universal values. The circular structure of values put forward by Schwartz depicts the relations between values (see **Figure 1**). In the circle, congruent values are adjacent to one another, whereas conflicting values are in opposing directions from the center. The values have bipolar dimensions: (1) self-enhancement versus self-transcendence and (2) openness to change versus

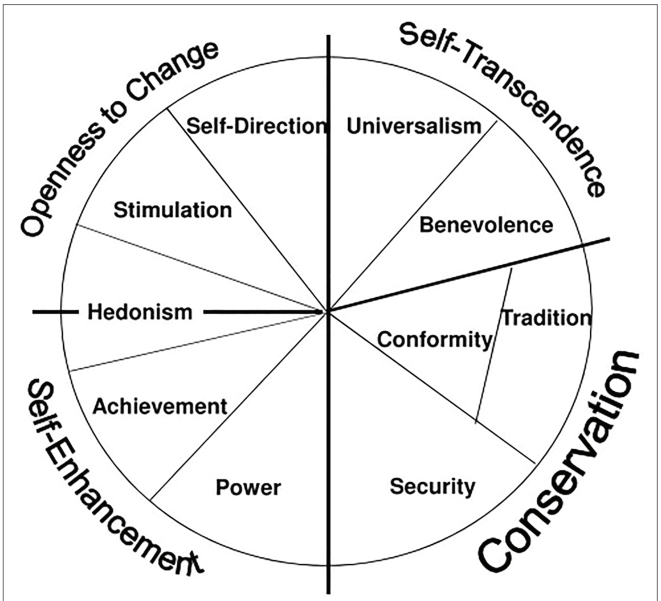


FIGURE 1 | The Circular Structure of Values Relations and Inter-Relatedness. This figure represents the theoretical model of relations among ten types of value. From Schwartz (2012b). The figure has been used with the permission of the author.

conservation. Self-enhancement values (power, achievement) emphasize and legitimize the pursuit of one's own interests, whereas self-transcendence values (universalism, benevolence) legitimize the pursue of the welfare of others. Openness to change values (self-direction, stimulation) encourage change, new ideas, and experiences, whereas conservation values (security, tradition, conformity) emphasize maintaining the establishment of shared rules and avoiding threat. Hedonism values share elements of openness and self-enhancement. Values of self-transcendence and conservation are externally focused and represent the association with others. In contrast, values of self-enhancement and openness to change are self-focused and represent the expression of personal interests and characteristics (Schwartz, 2012b). In fact, these ten values have been split into finer subparts which construct 19 values (Schwartz, 2017). However, these subparts are not included in this paper for the sake of simplicity and because these ten values already construct a broad and comprehensive understanding of values.

The value theory of Schwartz suggests that intra-individual value importance change occurs in a coherent manner (Schwartz, 1992). This organized pattern argues that if one value increases in importance while the opposing value remains stable, it may cause uneasy feelings, inner dissonance due to the inconsistency (Daniel and Benish-Weisman, 2019). Therefore, it is expected that adjacent values change in the same direction and values on the opposing side of the circle simultaneously change in the opposite direction (Bardi et al., 2009). Bardi et al. (2009) examined intra-individual change in value structures among adolescents longitudinally. They found that increases in the importance of each value were accompanied by those of compatible values and decreases in the importance of opposing

values. Specifically, adolescents who wish to stand on their own, which corresponds to the value of openness to change, feel inner-conflict because of their simultaneous desire to conform to their peer group, which corresponds to the value of conservation. In a similar manner, adolescents who strive for dominance in their group, which corresponds to the value of self-enhancement, may feel less inclined to tolerate variety among peers, which corresponds to self-transcendence. Such distress and conflict lead to changes in values hierarchies. In an experimental study, it was shown that increases in the manipulated values were accompanied by decreases in opposing values that were not manipulated (Maio et al., 2009).

Apart from inner value conflict, contradicting values occur interpersonally, too. For example, students and teachers might have opposing values, which leads to emotional and behavioral consequences among them. The emotional and behavioral consequences are the result of student and teacher perceptions. In other words, interpersonal value synergy or a "click" between teachers and students potentially depends on value congruence. In order to understand how a teacher and a student can be a good match, it is necessary to consider the value congruence.

Value Congruence

Individuals behave congruently with social events and social environments that are in line with their personal values. Lack of alignments or conflict leads to adverse attitudinal and behavioral expressions between individual and/or collective value orientations (Vauclair and Fischer, 2011). Person-environment fit (P-E fit) in the job setting can be "broadly defined as the compatibility between an individual and a work environment that occurs when their characteristics are well matched" (Kristof-brown et al., 2005, p. 281). In a broader sense, it can be defined as the match, congruence, similarity, or correspondence between the person and the environment. The mechanisms of person-environment fit function through the process of need fulfillment (Rice et al., 1989; Van Vianen, 2000). Achieving person-environment fit is a way to have individual needs met. The need fulfillment processes which are related to experiencing person-environment fit suggest that if needs are satisfied, individuals will experience more positive attitudes. On this basis, theories, such as the social comparison theory (Festinger, 1954), the balance theory (Heider, 1958), the similarity-attraction paradigm (Byrne, 1971), and the attraction-selection-attrition framework (Schneider, 1987), suggest that people have a fundamental need for consensual validation of their perspectives, which can be met by interacting with similar others. For example, individuals are attracted to careers that match with their values, therefore in the situation of a mismatch dropout rates increase (Holland, 1985, 1997; Schneider et al., 1995; Knafo and Sagiv, 2004). Consequently, people in a congruent environment will develop or benefit more than those in an incongruent environment (Feldman et al., 2001). When P-E fit is applied to the classroom and educational environment, students might have more benefits in a classroom that provides a congruent environment and from a teacher who performs the expected roles that fit the environment, that is, who provides a good quality teaching that helps students learn.

Researchers not only investigate the congruency of the environment but also the extent to which one's values are congruent with those that are salient in a particular context and among significant others, such as with a job manager or parents. Along these lines, the similarity-attraction paradigm shows that individuals are attracted by others who share similar characteristics (Byrne, 1971). Similarity binds people together because similar values facilitate communication and reduce uncertainty in interpersonal relationships (Kalliath et al., 1999; Edwards and Cable, 2009). Numerous studies have examined student-teacher congruence, particularly matches in student-teacher gender, race (Oates, 2003; Bates and Glick, 2013), thinking style (Zhang, 2006), beliefs (Polat, 2009), and expectations (Turner, 2009), and its relationship with learning achievement and test performance (e.g., Oates, 2003). The congruence in underlying characteristics (e.g., beliefs, expectations, values) might function differently from congruence in appearance (e.g., gender, race) because the congruence in appearance is more explicit and visible than the congruence in underlying innate characteristics. Westerman et al. (2002) illustrated that value congruence is a predictor of student satisfaction. Value congruence creates an environment in which individuals can freely express personal values and attain shared goals supported in the environment (Sagiv and Schwartz, 2000). This value match influences students' academic achievement because values shape lives, influence actions, gives expression to underlying beliefs (Rokeach, 1973); thus, it reflects to what extent students feel close to their teachers and how much students like their teachers. Given the value congruence, situations allowing for opposing student-teacher values may lead to more variations in student perceptions compared to situations that nurture value congruence.

Cultural Values

Learning culture represents a set of shared beliefs, values, and attitudes favorable to learning. Students' learning is influenced by learning culture at the school level. Schools' learning culture is influenced by higher cultural contexts at the regional and national levels. In this respect, cultures operate to shape the environment, behavior, and minds of their members. Value orientations at the cultural level provide a higher-order frame for understanding differences in student perceptions across cultural settings. In this paper, Schwartz's (1994, 1999, 2006) cultural-level value orientations are consulted because it is acknowledged that cultural-level processes are distinct from individual-level processes. These cultural-level value orientations are considered as "the cultures of the national groups" (Schwartz, 1999, p. 25).

At the culture level, seven cultural value orientations can be captured along three distinct dimensions (Schwartz, 1994, 1999, 2006). Schwartz argues that there are three basic issues in every society: (1) to what extent an individual belongs to a group, (2) how to preserve the social fabric, and (3) how to relate to the social world. As seen in **Figure 2**, bipolar

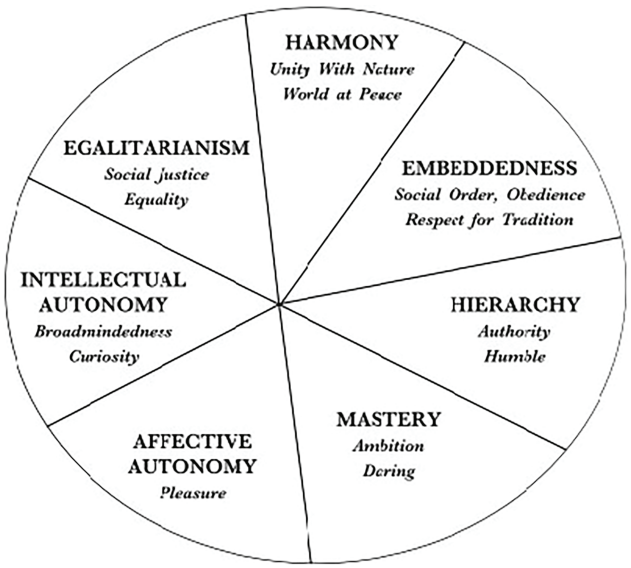


FIGURE 2 | Structure of Culture-Level Value Types. This figure represents the theoretical model of relations among ten types of value at the cultural level. From Schwartz (2006). The figure has been used with the permission of the author.

value orientations are formed, and cultural influences on which pole of the orientation are emphasized in society. Regarding the first societal issue, autonomy vs. embeddedness, autonomy partitions into affective autonomy and intellectual autonomy, which refers to values that encourage people to pursue positive experiences for themselves as well as their own ideas and intellectual aspirations. In autonomous cultures, individuals are encouraged to think, feel, and act as unique persons. The opposite pole, embeddedness, encourages traditional order, shared goals, ways of living, and the maintenance of the status quo is seen as a priority. The second societal issue produces the value types of hierarchy vs. egalitarianism. The hierarchy represents the hierarchical social order and unequal resource allocation; in contrast, egalitarianism emphasizes equality, mutual concern, and cooperation for everyone's welfare. The last societal problem produces the values of harmony vs. mastery. The former reflect societal discourses in which the social and natural world is acknowledged as it is, and fitting in harmoniously is emphasized. Individuals strive for a world at peace and the protection of the environment. The latter depicts the active control of the social and natural environment through self-assertion values. Individuals are ambitious, seek success, and competence in order to attain group or personal goals.

Similar to the individual level of value orientations, poles of the cultural value orientation tend to conflict with each other. Prevailing cultural value orientations are assumed as ideals that promote coherence, whereas incompatible values are likely to generate tension, elicit criticism, and pressure to change (Schwartz, 2006). Examples thereof are examined in the frame of prejudice and aggression toward an outgroup with different value orientations (Struch and Schwartz, 1989). In a society where hierarchical relationships are legitimate, individuals tend to mark off in and outgroups. In contrast, in cultures higher on egalitarianism, emphasizing the equality of all people in the world, members of outgroups are less likely to be devalued (Tajfel and Turner, 1979; Schiefer et al., 2010).

Applied to the educational setting (see **Table 1**), it is hypothesized that in certain cultures where, for example, harmony is emphasized it is more probable that students and teachers will share the same harmonious values to a greater extent compared to the opposite pole. This will lead to less openness to communicate about differences. Therefore, in these communities where more alignment occurs on harmony, we expect less variety in student perceptions with regard to teaching quality. Similarly, in certain cultures where, for example, embeddedness is emphasized, less variety in student perceptions with regard to teaching quality is expected because in these cultures individuals are viewed as entities embedded in the collectivity with shared goals. Along the same line, if hierarchy is emphasized in a culture, it is more probable that students will consider teachers at a higher-order level and feel responsible to obey. Therefore, it is more probable that they will share similar opinions about teachers and extreme opinions will be less encouraged compared to the opposite pole. This will again lead to less variety in student perceptions with regard to teaching quality. Based on the studies in which the groupings of nations were depicted (see **Figure 3**; Schwartz, 1999, 2006), less student perceptions variety in, for example, Slovenia, Singapore, Indonesia, and South Korea, is expected because these cultures have adopted the pole of harmony, embeddedness, and hierarchy.

In contrast, in certain cultures where, for example, mastery is prioritized, self-assertion is seen as individual ambition and competence. Therefore, because individual opinion differences are expected to be openly discussable, more variety in student perceptions is to be expected. Similarly, if autonomy is prioritized in a certain culture, an individual's unique thinking, feeling, and acting are encouraged, which is expected to lead to a variety of student perceptions. Along a similar line, in cultures where, for example, egalitarianism is prioritized, embracing different opinions and value orientations is more probable. In these environments, students and teachers are hypothesized not to hesitate to express own value orientations, leading to tolerance of diverse values orientations, and more variety in student perceptions. Based on groupings of nations (see **Figure 3**; Schwartz, 1999, 2006), more student perceptions variety in,

TABLE 1 | The relationship between cultural values, student perception, and social desirability.

Cultural values	Manifestations of student perception	Social desirability
Harmony Embeddedness Hierarchy	More congruence, less variability of perception	More impression management, less self-deceptive positivity
Mastery Affective autonomy Intellectual autonomy Egalitarianism	Less congruence, more variability of perception	Less impression management, more self-deceptive positivity

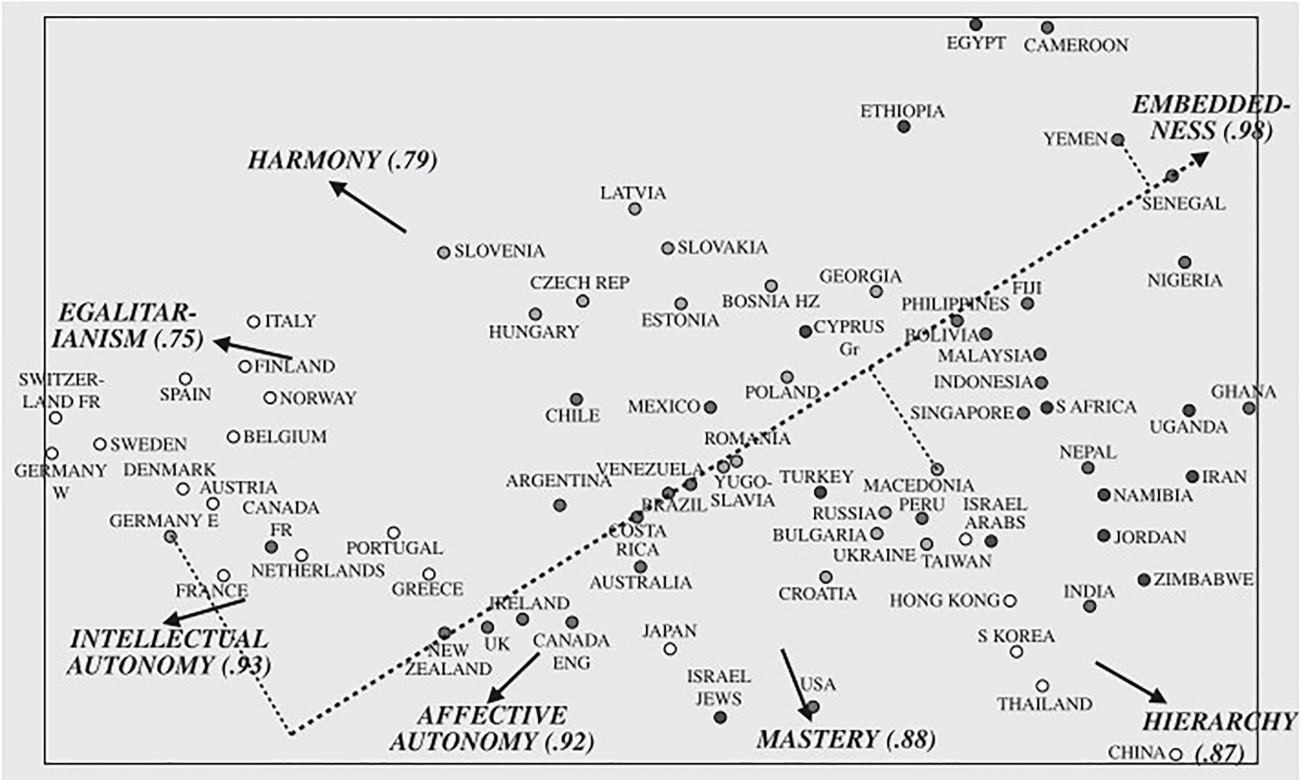


FIGURE 3 | Cultural Value Orientations from 76 National Groups. The figure represents cultural value orientations from 76 national groups based on the combined teacher and college student samples. From Schwartz (2006). The figure has been used with the permission of the author.

for example, the United States, Canada, The Netherlands, and Spain, is expected because these cultures have adopted the pole of mastery, autonomy, and egalitarianism values.

One of the important features of cultural value orientations is that cultural values shape and justify individual and group beliefs, actions, and goals (Schwartz, 2006). In every layer of society, from child-rearing and everyday practices to institutional regulations and policies, there are traces of cultural values emphasized in the society. Indeed, it is inevitable to expect the projections of cultural value orientations in the classroom and educational settings. For example, schools in primary and secondary education in The Netherlands need to pay attention to citizenship education. This education ensures that students develop social and societal competencies and respect for differences in every aspect such as religions, beliefs, political views, and sexual orientations. Regardless of these differences, schools need to ensure an environment in which students and staff feel safe and accepted. Students learn to listen to each other, determine their opinion and express it constructively, as well as, notice and respect different opinions. These education practices coincide with the emphasis of the cultural orientation, which is hypothesized to be at the pole of mastery, autonomy, and egalitarianism value for The Netherlands (Schwartz, 2006, 2014). It is therefore important to recognize the nested nature of education.

Another important feature is that cultural value orientations can change, yet the change is slow (Schwartz et al., 2000). Although cultural orientations can persist over hundreds of years, they change gradually. The augmentation of wealth and contact with other cultures by globalization, the massive effect of the pandemic, and the advancement in technology, all together may lead to change in cultural value orientations. Due to this slow and gradual change, a change in educational practices and policies is also to be expected.

SOCIAL DESIRABILITY

Understanding how socially desirable responding (SDR) impacts student self-reporting practices might be useful to understand the influence of shared values on student perceptions (Schwartz et al., 1997). SDR (Crowne and Marlowe, 1960) refers to the tendency to respond in a manner that strengthens social approval instead of reflecting one's true feelings. Concerns of social desirability may be inevitable in the case of values because values are the goals and preferences individuals consider socially desirable. Respondents are likely to respond in a biased manner to the degree that certain values are strongly encouraged within the social environment in certain social contexts (Fisher and Katz, 2000). In contrast, values that are of marginal importance are less likely to influence responses, meaning that the motivation of social desirability bias will be weak.

Paulhus (1991) divides two factors of social-desirability bias. The first factor, *impression management*, is the desire to represent oneself in a socially conventional way. Individuals who have higher scores on impression management tend to be more sensitive to social influence. It was found that impression management is more related to values highlighting the importance of social harmony (i.e., conservation and self-transcendence) rather than to those characterized by a personal focus (i.e., openness to change and self-enhancement; Danioni and Barni, 2020). The second factor, *self-deceptive positivity*, reflects a favorable self-presentation. Individuals who score higher on self-deceptive positivity hold values characterized by a personal focus. A positive link between self-deceptive positivity and self-enhancement values and a lack of relationship between self-deceptive positivity with conservation and self-transcendence demonstrates that self-deceptive positivity is characterized by a focus on the self (Danioni and Barni, 2020).

For an individual to indicate SDR, the individual needs to have some knowledge of what would be desirable in the corresponding cultural context (i.e., social-norm intelligence; Bou Malham and Saucier, 2016). Hence, SDR involves implicit reference to culturally shared norms, standards, and values. Although evidence is mixed as to whether individuals in different cultures exhibit SDR to the same extent, much research supports cross-cultural differences. For example, it was argued that SDR can be an adaptive response strategy in certain situations, in which strong norms are shared around a particular issue and are learned through socialization (Ross and Mirowsky, 1984). Research shows that individuals in collectivistic cultures tend to have higher impression management and lower self-deceptive scores, whereas individuals in individualistic cultures tend to have higher self-deceptive scores (Lalwani et al., 2006). For example, it was shown that Hong Kong participants scored higher on impression management than did US participants, whereas US participants scored higher on self-deceptive than did Hong Kong participants (Lalwani et al., 2009). Such findings suggest that impression management and self-deceptive positivity reflect different cultural settings. Given the relationship between SDR and cultural contexts, SDR can be interpreted as cultural consonance to maintain person-environment or person-group congruence (Bou Malham and Saucier, 2016).

In the school context, peers play a decisive and critical role in students' behaviors and attitudes. Through friendship, peers influence student's academic functioning (Newcomb and Bagwell, 1995; Rambaran et al., 2017). This academic functioning includes involvement in school (Kindermann, 2007), motivation (Wentzel et al., 2010; Molloy et al., 2011), and reading achievements (Cooc and Kim, 2016). As seen in **Table 1**, it would be expected that the variety of student perceptions changes due to the influence of peers concerning SDR depending on cultural contexts. Specifically, in certain cultures where the pole of harmony, embeddedness, and hierarchy is emphasized, it is hypothesized that students are more likely to engage in impression management and less likely in self-deceptive positivity to maintain good relationships with others compared to the opposite pole. In these cultural settings, high impression management would likely serve the function of group harmony, which leads to more congruence and less student perceptions variability. In contrast, in certain cultures where the pole of mastery, autonomy, and egalitarianism is emphasized, it is less likely that students engage in impression management and more likely to engage in self-deceptive positivity. In these cultural settings, because self-deceptive positivity is characterized by a focus on the self, individuals are likely to reveal their self-identity by having and expressing their unique and distinctive opinions, which leads to less congruence and more variability among student perceptions.

DISCUSSION AND CONCLUDING REMARKS

This conceptual paper explores the theoretical underpinning of value orientations as antecedents of the variation of student perceptions within and across countries and aims to shed light on how value orientations and social desirability may explain variability in student perceptions. In other words, it aims to provide insights into understanding value orientations at the individual and collective level to better understand manifestations of student perceptions. The added value of Schwartz's value theory promotes understanding of student perceptions through value orientations. Student evaluation of teaching quality is treated as an instance of perceptions, and value orientations as underlying reasoning behind perceptions.

As seen in the proposed conceptual model (see **Figure 4**), it is postulated that value orientations, value congruence at the individual- and cultural-level, and social desirability may play a role in student perceptions of teaching quality. According to the model, (cultural) value orientations may play a role in influencing every layer of society, including social interactions manifested as value congruence and social desirability. In this sense, teacher's and student's individual value orientations could be in interaction with cultural value orientations, which in turn generate value (dis)congruence. Similarly, cultural value orientations could be related to two factors of social desirability, namely, impression management and self-deceptive positivity. Taken together, manifestations of student perception might be influenced by the specific cultural value orientation, value congruence between teacher and student, and social desirability (see **Table 1**). This model can serve as the theoretical rationale guiding future empirical research for investigating the antecedents of student perception variability of teaching quality. Also, the model can be used to understand the universal reasoning behind the empirical research that investigates the consequences of student perception variability (e.g., Schweig, 2016; Schenke et al., 2017; Bardach et al., 2019a). Empirical research reported that lower levels of student variability are associated with better learning outcomes. For example, it was found that classrooms with higher levels of consensus (i.e., lower levels of variability) regarding the teacher's ability to effectively control the classroom experienced more effective instruction and vice versa (Schweig, 2016). The model can serve to support teachers toward higher levels of value congruence to enhance a safe and stimulating classroom climate that is fundamental for learning.

This conceptual framework applied to educational settings should be investigated empirically in future research. Besides,

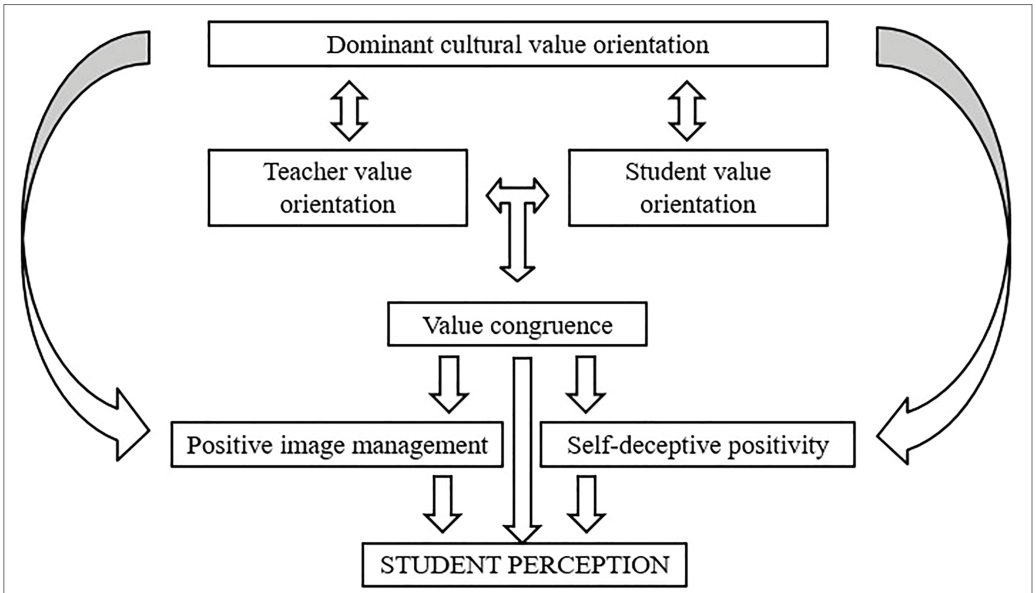


FIGURE 4 | The Proposed Conceptual Framework.

individual factors (e.g., age and gender), class context (e.g., learning environment), and school context (e.g., private vs. public) need to be taken into consideration even though they are not portrayed in the conceptual framework. In order to explain the significance of these factors, student age in connection with the development of values can be given as an illustration. With a cross-sectional study, it was revealed that adolescents attribute lower importance to other-focused values and higher importance to self-focused values compared to adults, and values that emphasize autonomy and self-direction peak in late-adolescence (Schwartz, 2012a). Longitudinal studies depict similar findings, regarding students' values change across the school year (Hofmann-Towfigh, 2007). It was found that adolescents' self-focused values of self-direction, power, and achievement increase while the other-focused values of self-transcendence, benevolence, and universalism decrease over time. This trend is connected to the natural phenomenon that adolescence is a period of exploration, independence, and identity development. In connection with these developmental processes, it would be expected that young students will be more likely to be influenced by others, that is, peers, because they do not form a self-construal yet. Consequently, it is expected that those students are less likely to have unique opinions on teaching quality, which leads to less variety of student perceptions. In contrast, older students who are at the later stages of their identity and self-construal development are more likely to engage in agentic behavior, which leads to more variety of student perceptions with regard to teaching quality.

From a practical point of view, this paper emphasizes the importance of teachers being aware of and acknowledging different value orientations in the classroom. Students have a personal value orientation, which makes each student unique and special. If teachers assume that students hold the same values with each other and with the teacher, it becomes difficult to create a safe learning environment that supports trust, collaboration, and transparent sharing. At this point, teachers need to be supported to acknowledge individual and cultural diversities. Indeed, teacher education training could incorporate these aspects into their programs, and professional development initiatives could be developed to assist the development of experienced teachers in this respect.

For the sake of educational research and policy-makers, it is also important to emphasize that the extent to which students vary in their perceptions of teaching can be considered a rich source of information when considered from a bigger picture that incorporates cultural orientations. Less variability of student perceptions does not necessarily mean that student perceptions are less valuable, less informative, or less reliable. It simply means that teachers, educational researchers, and policy-makers need to be aware of the influence of shared values, impression management, and self-deceptive positivity reflected in student perceptions of teaching quality. Whether the variability of student perceptions is low or high, students have conceptions about how teaching quality may help or hinder their learning. Therefore, it is essential to emphasize the significance of student perceptions *per se*.

AUTHOR CONTRIBUTIONS

Each author has made a substantial contribution to the conception of the work, substantively revised it, and approved the submitted version. BG, MH-L, RM, and EJ: conceptualization. BG: writing – original draft preparation. MH-L, RM, and EJ: writing – review and editing and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This study was part of the Ph.D. project of the first author.

REFERENCES

André, S., Maulana, R., Helms-Lorenz, M., Telli, S., Chun, S., Fernández-García, C.-M., et al. (2020). Student perceptions in measuring teaching behavior across six countries: A multi-group confirmatory factor analysis approach to measurement invariance. *Front. Psychol.* 11, 1–19. doi: 10.3389/fpsyg.2020.00273

Banks, T., and Zionts, P. (2009). Teaching a cognitive behavioral strategy to manage emotions rational emotive behavior therapy in an educational setting. *Interv. Sch. Clin.* 44, 307–313. doi: 10.1177/1053451208330893

Bardach, L., Lüftenegger, M., Yanagida, T., Schober, B., and Spiel, C. (2019a). The role of within-class consensus on mastery goal structures in predicting socio-emotional outcomes. *Br. J. Educ. Psychol.* 89, 239–258. doi: 10.1111/bjep.12237

Bardach, L., Lüftenegger, M., Yanagida, T., Spiel, C., and Schober, B. (2019b). Achievement or agreement – which comes first? Clarifying the temporal ordering of achievement and within-class consensus on classroom goal structures. *Learn. Instr.* 61, 72–83. doi: 10.1016/j.learninstruc.2019.01.003

Bardach, L., Yanagida, T., Morin, A. J. S., and Lüftenegger, M. (2021). Is everyone in class in agreement and why (not)? Using student and teacher reports to predict within-class consensus on goal structures. *Learn. Instr.* 71:101400. doi: 10.1016/j.learninstruc.2020.101400

Bardi, A., Lee, J. A., Hofmann-Towfigh, N., and Soutar, G. (2009). The structure of intraindividual value change. *J. Pers. Soc. Psychol.* 97, 913–929. doi: 10.1037/a0016617

Bates, L. A., and Glick, J. E. (2013). Does it matter if teachers and schools match the student? Racial and ethnic disparities in problem behaviors. *Soc. Sci. Res.* 42, 1180–1190. doi: 10.1016/j.ssresearch.2013.04.005

Bernard, M. E. (1990). Rational-emotive therapy with children and adolescents: treatment strategies. *Sch. Psychol. Rev.* 19, 294–303. doi: 10.1080/02796015.1990.12085468

Bou Malham, P., and Saucier, G. (2016). The conceptual link between social desirability and cultural normativity. *Int. J. Psychol.* 51, 474–480. doi: 10.1002/ijop.12261

Byrne, D. (1971). *The Attraction Paradigm*. New York: Academic Press.

Clawson, J. G. (2012). *Level Three Leadership: Getting below the Surface. 5th Edn*. New Jersey: Pearson Education.

Cooc, N., and Kim, J. S. (2016). Peer influence on children's reading skills: A social network analysis of elementary school classrooms. *J. Educ. Psychol.* 109, 727–740. doi: 10.1037/edu0000166

Creemers, B. P. M. (1994). *The Effective Classroom*. London: Cassell.

Creemers, B. P. M., and Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. United Kingdom: Routledge.

Crowne, D. P., and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.* 24, 349–354. doi: 10.1037/h0047358

Daniel, E., and Benish-Weisman, M. (2019). Value development during adolescence: dimensions of change and stability. *J. Pers.* 87, 620–632. doi: 10.1111/jopy.12420

Danielson, C. (2013). *The Framework for Teaching: Evaluation Instrument*. Princeton, NJ: Danielson Group.

Danioni, F., and Barni, D. (2020). Value priorities, impression management and self-deceptive enhancement: once again, much substance and a little bit of style. *J. Soc. Psychol.* 161, 146–159. doi: 10.1080/00224545.2020.1778619

den Brok, P. J. (2001). *Teaching and Student Outcomes: A Study on Teachers' Thoughts and Actions from an Interpersonal and a Learning Activities Perspective*. United States: WCC Publications.

Edwards, J. R., and Cable, D. M. (2009). The value of value congruence. *J. Appl. Psychol.* 94, 654–677. doi: 10.1037/a0014891

Ellis, A. (1962). *Reason and Emotion in Psychotherapy*. New York: L. Stuart.

Ellis, A. (1974). Experience and rationality: The making of a rational-emotive therapist. *Psycho. Theory Res. Practice* 11, 194–198. doi: 10.1037/h0086338

Ellis, A. (1991). The ABC's of RET. *Humanist* 3, 11–12.

Feldman, K. A., Ethington, C. A., and Smart, J. C. (2001). A further investigation of major field and person-environment fit: sociological versus psychological interpretations of Holland's theory. *J. High. Educ.* 72, 670–698. doi: 10.1080/00221546.2001.11777121

Festinger, L. (1954). A theory of social comparison processes. *Hum. Relat.* 7, 117–140. doi: 10.1177/001872675400700202

Fischer, R., Vauclair, C. M., Fontaine, J. R. J., and Schwartz, S. H. (2010). Are individual-level and country-level value structures different? Testing Hofstede's legacy with the Schwartz value survey. *J. Cross-Cult. Psychol.* 41, 135–151. doi: 10.1177/0022022109354377

Fisher, R. J., and Katz, J. E. (2000). Social-desirability bias and the validity of self-reported values. *Psychol. Mark.* 17, 105–120. doi: 10.1002/(SICI)1520-6793(200002)17:2<105::AID-MAR3>3.0.CO;2-9

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: testing a developmental framework of teaching effectiveness in over 4,000 classrooms. *Elem. Sch. J.* 113, 461–487. doi: 10.1086/669616

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. United Kingdom: Taylor & Francis.

Heider, F. (1958). *The Psychology of Interpersonal Relations*. United States: John Wiley and Sons Inc.

Helms-Lorenz, M., and Visscher, A. J. (2021). Unravelling the challenges of the data-based approach to teaching improvement. *Sch. Eff. Sch. Improv.*, 1–23. doi: 10.1080/09243453.2021.1946568

Hofmann-Towfigh, N. (2007). Do students' values change in different types of schools? *J. Moral Educ.* 36, 453–473. doi: 10.1080/03057240701688010

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA: SAGE.

Holland, J. L. (1985). *Making Vocational Choices: A Theory of Careers*. Englewood Cliffs, NJ: Prentice-Hall.

Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments. 3rd Edn*. Boston: Psychological Assessment Resources INC.

Inda-Caro, M., Maulana, R., Fernández-García, C. M., Peña-Calvo, J. V., del Rodríguez-Menéndez, M. C., and Helms-Lorenz, M. (2019). Validating a model of effective teaching behaviour and student engagement: perspectives from Spanish students. *Learn. Environ. Res.* 22, 229–251. doi: 10.1007/s10984-018-9275-z

Kalliath, T. J., Bluedorn, A. C., and Strube, M. J. (1999). A test of value congruence effects. *J. Organ. Behav.* 20, 1175–1198. doi: 10.1002/(SICI)1099-1379(199912)20:7<1175::AID-JOB960>3.0.CO;2-5

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Dev.* 78, 1186–1203. doi: 10.1111/j.1467-8624.2007.01060.x

Knafo, A., and Sagiv, L. (2004). Values and work environment: mapping 32 occupations. *Eur. J. Psychol. Educ.* 19, 255–273. doi: 10.1007/BF03173223

Kristof-brown, A. L., Zimmerman, R. D., and Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Pers. Psychol.* 58, 281–342. doi: 10.1111/j.1744-6570.2005.00672.x

Lalwani, A. K., Shavitt, S., and Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *J. Pers. Soc. Psychol.* 90, 165–178. doi: 10.1037/0022-3514.90.1.165

Lalwani, A. K., Shrum, L. J., and Chiu, C. (2009). Motivated response styles: The role of cultural values, regulatory focus, and self-consciousness in socially desirable responding. *J. Pers. Soc. Psychol.* 96, 870–882. doi: 10.1037/a0014622

Lüdtke, O., Trautwein, U., Kunter, M., and Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment: A reanalysis of TIMSS data. *Learn. Environ. Res.* 9, 215–230. doi: 10.1007/s10984-006-9014-8

Maio, G. R., Pakizeh, A., Cheung, W.-Y., and Rees, K. J. (2009). Changing, priming, and acting on values: effects via motivational relations in a circular model. *J. Pers. Soc. Psychol.* 97, 699–715. doi: 10.1037/a0016420

Maulana, R., and Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: construct representation and predictive quality. *Learn. Environ. Res.* 19, 335–357. doi: 10.1007/s10984-016-9215-8

Maulana, R., Helms-Lorenz, M., and van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teach. Teach. Educ.* 51, 225–245. doi: 10.1016/j.tate.2015.07.003

Maulana, R., Opdenakker, M. C., and Bosker, R. (2016). Teachers' instructional behaviors as important predictors of academic motivation: changes and links

across the school year. *Learn. Individ. Differ.* 50, 147–156. doi: 10.1016/j.lindif.2016.07.019

Meade, A. W., and Eby, L. T. (2007). Using indices of group agreement in multilevel construct validation. *Organ. Res. Methods* 10, 75–96. doi: 10.1177/1094428106289390

Molloy, L. E., Gest, S. D., and Rulison, K. L. (2011). Peer influences on academic motivation: exploring multiple methods of assessing youths' most "influential" peer relationships. *J. Early Adolesc.* 31, 13–40. doi: 10.1177/0272431610384487

Newcomb, A. F., and Bagwell, C. L. (1995). Children's friendship relations: A meta-analytic review. *Psychol. Bull.* 117, 306–347. doi: 10.1037/0033-2909.117.2.306

Oates, G. L. S. C. (2003). Teacher-student racial congruence, teacher perceptions, and test performance. *Soc. Sci. Q.* 84, 508–525. doi: 10.1111/1540-6237.8403002

Parashar, S., Dhar, S., and Dhar, U. (2004). Perception of values: A study of future professionals. *J. Hum. Values* 10, 143–152. doi: 10.1177/097168580401000207

Paulhus, D. L. (1991). "Measurement and control of response bias," in *Measures of Personality and Social Psychological Attitudes*. eds. J. P. Robinson, P. Shaver and L. S. Wrightsman (San Diego: Academic Press).

Pianta, R. C., and Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: standardized observation can leverage capacity. *Educ. Res.* 38, 109–119. doi: 10.3102/0013189X09332374

Pianta, R. C., La Paro, K., and Hamre, B. K. (2008). *Classroom Assessment Scoring System*. United States: Paul H Brookes Publishing.

Polat, N. (2009). Matches in beliefs between teachers and students, and success in L2 attainment: The Georgian example. *Foreign Lang. Ann.* 42, 229–249. doi: 10.1111/j.1944-9720.2009.01019.x

Rambaran, J. A., Schwartz, D., Badaly, D., Hopmeyer, A., Steglich, C., and Veenstra, R. (2017). Academic functioning and peer influences: A short-term longitudinal study of network-behavior dynamics in middle adolescence. *Child Dev.* 88, 523–543. doi: 10.1111/cdev.12611

Rice, R. W., McFarlin, D. B., and Bennett, D. E. (1989). Standards of comparison and job satisfaction. *J. Appl. Psychol.* 74, 591–598. doi: 10.1037/0021-9010.74.4.591

Roccas, S., Sagiv, L., Schwartz, S. H., and Knafo, A. (2002). The big five personality factors and personal values. *Personal. Soc. Psychol. Bull.* 28, 789–801. doi: 10.1177/0146167202289008

Rokeach, M. (1973). *The Nature of Human Values*. New York: Free Press.

Ross, C. E., and Mirowsky, J. (1984). Socially-desirable response and acquiescence in a cross-cultural survey of mental health. *J. Health Soc. Behav.* 25, 189. doi: 10.2307/2136668

Sagiv, L. (2002). Vocational interests and basic values. *J. Career Assess.* 10, 233–257. doi: 10.1177/1069072702010002007

Sagiv, L., and Schwartz, S. H. (2000). Value priorities and subjective well-being: direct relations and congruity effects. *Eur. J. Soc. Psychol.* 30, 177–198. doi: 10.1002/(SICI)1099-0992(200003/04)30:2<177::AID-EJSP982>3.0.CO;2-Z

Sammons, P., and Bakkum, L. (2011). Effective schools, equity and teacher effectiveness: A review to the literature. *Prof. Rev. Curr. Form. Del. Prof.* 15, 9–26.

Scheerens, J. (2016). *Educational Effectiveness and Ineffectiveness: A Critical Review of the Knowledge Base*. Germany: Springer.

Scheerens, J., Luyten, H., Steen, R., and Luyten-de Thouars, Y. (2007). Review and meta-analyses of school and teaching effectiveness. Available at: https://www.utwente.nl/nl/bms/omd/Medewerkers/artikelen/scheerens/rev-metaanalysis2007.PDF (Accessed July 31, 2019).

Schenke, K., Ruzek, E., Lam, A. C., Karabenick, S. A., and Eccles, J. S. (2017). Heterogeneity of student perceptions of the classroom climate: a latent profile approach. *Learn. Environ. Res.* 20, 289–306. doi: 10.1007/s10984-017-9235-z

Schenke, K., Ruzek, E., Lam, A. C., Karabenick, S. A., and Eccles, J. S. (2018). To the means and beyond: understanding variation in students' perceptions of teacher emotional support. *Learn. Instr.* 55, 13–21. doi: 10.1016/j.learninstruc.2018.02.003

Schiefer, D., Möllering, A., Daniel, E., Benish-Weisman, M., and Boehnke, K. (2010). Cultural values and outgroup negativity: A cross-cultural analysis of early and late adolescents. *Eur. J. Soc. Psychol.* 40, 635–651. doi: 10.1002/ejsp.745

Schneider, B. (1987). The people make the place. *Pers. Psychol.* 40, 437–453. doi: 10.1111/j.1744-6570.1987.tb00609.x

Schneider, B., Goldstein, H. W., and Smith, B. D. (1995). The Asa framework: An update. *Pers. Psychol.* 48, 747–773. doi: 10.1111/j.1744-6570.1995.tb01780.x

Schwartz, S. H. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Adv. Exp. Soc. Psychol.* 25, 1–65. doi: 10.1016/S0065-2601(08)60281-6

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *J. Soc. Issues* 50, 19–45. doi: 10.1111/j.1540-4560.1994.tb01196.x

Schwartz, S. H. (1999). A theory of cultural values and some implications for work. *Appl. Psychol. Int. Rev.* 48, 23–47. doi: 10.1111/j.1464-0597.1999.tb00047.x

Schwartz, S. H. (2006). A theory of cultural value orientations: explication and applications. *Comp. Sociol.* 5, 137–182. doi: 10.1163/ej.9789004158207.i-193.19

Schwartz, S. H. (2012a). "Values and religion in adolescent development: cross-national and comparative evidence," in *Values, Religion, and Culture in Adolescent Development*. eds. G. Trommsdorff and X. Chen (Cambridge, England: Cambridge University Press), 97–122.

Schwartz, S. H. (2012b). An overview of the Schwartz theory of basic values. *Online Readings Psychol. Culture* 2, 2307–2919. doi: 10.9707/2307-0919.1116

Schwartz, S. H. (2014). "National culture as value orientations: consequences of value differences and cultural distance," in *Handbook of the Economics of Art and Culture. Vol. 2*. eds. V. A. Ginsburgh and D. Throsby (Netherlands: Elsevier), 547–586.

Schwartz, S. H. (2017). "The refined theory of basic values" in *Values and Behavior: Taking a Cross Cultural Perspective*. eds. S. Roccas and L. Sagiv (New York: Springer International Publishing), 51–72.

Schwartz, S. H., Bardi, A., and Bianchi, G. (2000). "Value adaptation to the imposition and collapse of communist regimes in Eastern Europe," in *Political Psychology: Cultural and Cross Cultural Perspectives*. eds. S. A. Renshon and J. Duckitt (United Kingdom: Macmillan), 217–237.

Schwartz, S. H., Verkasalo, M., Antonovsky, A., and Sagiv, L. (1997). Value priorities and social desirability: much substance, some style. *Br. J. Soc. Psychol.* 36, 3–18. doi: 10.1111/j.2044-8309.1997.tb01115.x

Schweig, J. D. (2016). Moving beyond means: revealing features of the learning environment by investigating the consensus among student ratings. *Learn. Environ. Res.* 19, 441–462. doi: 10.1007/s10984-016-9216-7

Smith, T. W. (1982). Irrational beliefs in the cause and treatment of emotional distress: A critical review of the rational-emotive model. *Clin. Psychol. Rev.* 2, 505–522. doi: 10.1016/0272-7358(82)90027-7

Stern, P. C. (2000). Toward a coherent theory of environmentally significant behavior. *Journal of Social Issues.* 56, 407–424. doi: 10.1111/0022-4537.00175

Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., and Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Human Ecology Review.* 6, 81–97.

Struch, N., and Schwartz, S. H. (1989). Intergroup aggression: its predictors and distinctness from in-group bias. *J. Pers. Soc. Psychol.* 56, 364–373. doi: 10.1037/0022-3514.56.3.364

Tajfel, H., and Turner, J. (1979). "An integrative theory of intergroup conflict," in *The Social Psychology of Intergroup Relations*. eds. W. G. Austin and S. Worchel (Brooks/Cole: Monterey, CA)

Turner, M. (2009). Adult south Sudanese students in Australia: The significance of congruence between student and teacher expectations. *Prospect* 24, 48–59.

van de Grift, W. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educ. Res.* 49, 127–152. doi: 10.1080/00131880701369651

van Tartwijk, J., Brekelmans, M., Wubbels, T., Fisher, D. L., and Fraser, B. J. (1998). Students' perceptions of teacher interpersonal style: The front of the classroom as the teacher's stage. *Teach. Teach. Educ.* 14, 607–617. doi: 10.1016/S0742-051X(98)00011-0

Van Vianen, A. E. M. (2000). Person-organization fit: The match between newcomers' and recruiters' preferences for organizational cultures. *Pers. Psychol.* 53, 113–149. doi: 10.1111/j.1744-6570.2000.tb00196.x

Vauclair, C. M., and Fischer, R. (2011). Do cultural values predict individuals' moral attitudes? A cross-cultural multilevel approach. *Eur. J. Soc. Psychol.* 41, 645–657. doi: 10.1002/ejsp.794

Wentzel, K. R., Battle, A., Russell, S. L., and Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemp. Educ. Psychol.* 35, 193–202. doi: 10.1016/j.cedpsych.2010.03.002

Westerman, J. W., Nowicki, M. D., and Plante, D. (2002). Fit in the classroom: predictors of student performance and satisfaction in management education. *J. Manag. Educ.* 26, 5–18. doi: 10.1177/105256290202600102

Zhang, L. F. (2006). Does student-teacher thinking style match/mismatch matter in students' achievement? *Educ. Psychol.* 26, 395–409. doi: 10.1080/01443410500341262

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 October 2021
doi: 10.3389/fpsyg.2021.729360

Evaluating Educational Credentials of Teachers as Predictor of Effective Teaching: A Pupil Fixed-Effect Modeling Approach

*Ji Liu**

Faculty of Education, Shaanxi Normal University, Xi'an, China

Many factors serve as predictors of effective teaching; particularly, there is an ongoing debate regarding whether the educational credentials of teachers are indicative of their contribution to pupil's educational success. Utilizing pupil fixed-effect modeling and the China Education Panel Survey dataset (n = 5,032), this study evaluates the extent to which teachers who hold at least a Bachelor of Education (BEd) degree perform better than those who do not, in terms of pupil learning gains over the course of a full academic year. Empirical results from the pupil fixed-effect model indicate that mean learning gain is 0.042 SDs (95% CI: 0.008–0.083, p = 0.040) higher among pupils who studied with teachers holding higher educational credentials (at least a BEd degree) than those with lower educational credentials. This effect translates to approximately 1 month of additional learning per year, which is significant considering the potential compounding aggregation effects over the course of the entire educational career of pupils. This study adds new evidence that highlights the importance of the educational credentials of teachers as a predictor of effective teaching and that better-educated teachers can lead to improved pupil learning gains.

Keywords: China, learning gains, effective teachers, pupil fixed-effect, educational success

OPEN ACCESS

Edited by:
Ridwan Maulana,
University of Groningen, Netherlands

Reviewed by:
Emrah Soykan,
Near East University, Cyprus
Marcus Kiiti Borges,
University of Ponta Grossa (UEPG), Brazil

***Correspondence:**
Ji Liu
jiliu@snnu.edu.cn

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *23 June 2021*
Accepted: *23 September 2021*
Published: *26 October 2021*

Citation:
Liu J (2021) Evaluating Educational Credentials of Teachers as Predictor of Effective Teaching: A Pupil Fixed-Effect Modeling Approach. Front. Psychol. 12:729360. doi: 10.3389/fpsyg.2021.729360

INTRODUCTION

In recent decades, the quality of education has become a key topic of concern among parents, educators, and policymakers. The ongoing economic, social, and information integration holds strong implications for how the prospect of the future is changing, in that globalization, automation, and knowledge economy are driving a broad emphasis on investment in people. Conceptually, a wide range of inputs fed into the process of learning, such as pupil effort, teacher quality, and school facilities, while instructional and organizational practices are key determinants affecting how efficiently these inputs combine and generate meaningful learning engagements. To that end, evidence across international contexts underscores the critical importance of effective teaching for bolstering educational success, encouraging meaningful interaction, and sustaining lifelong learning (Altinok and Kingdon, 2012; Hanushek and Rivkin, 2012; Chetty et al., 2014). More critically, the weight of the influence of teachers on learning outcomes has been shown

to be more deterministic in developing countries, where availability of home educational resources and learning opportunities are scarce (Bau and Das, 2017; Liu and Steiner-Khamsi, 2020). Consequently, policymakers have become interested in evaluating how the educational credentials of teachers can serve as promising predictors of effective teaching and in understanding how such observable indicators of teacher quality affect the learning of children (Liu, 2021).

Nonetheless, the empirical literature on the value of the educational credentials of teachers has long been divided. In one camp, scholars argue that educational credential reflects positively on the level of the academic proficiency of a teacher, in addition to the reflections of sufficient cognitive and non-cognitive skills required to design, execute, and engage in teaching activities (Harris and Sass, 2011). The theoretical basis for this view rests on that the human capital development of teachers precedes in importance than that of their pupils, and that their educational mindset, instructional effectiveness, and professional learning all depend on a foundational level of accumulated skills throughout their own teacher education preparation. As a case in point, Jackson and Bruegmann (2009) summarized the attained level of observable human capital development of teachers, such as degree level, licensure and certification status, and score on licensure examination, into a single index and found positive associations with math and reading achievement of pupils. In broad strokes, teaching is a complex task that requires a well-rounded set of cognitive and non-cognitive abilities to excel and not to mention intensive demands for pedagogical skills and subject knowledge for classroom success. Studies supporting this view have persistently shown that the human capital attainment of teachers strongly predicts on-the-job performance and subsequent pupil learning gains in the classroom (Bastian, 2019; Noell et al., 2019).

Conversely, there is an opposing research cluster that casts doubt on the positive link between the educational credentials of teachers and the academic achievement of pupils. In particular, scholars argue that observable teacher traits are inconsistent at best when used as the predictors of effective teaching (Hanushek and Rivkin, 2006; Chingos and Peterson, 2011). These studies directly put in question the usefulness of existing teacher-education arrangements, as well as making uncertain the effectiveness of educational policy initiatives aimed at improving teacher educational attainment and teacher preparation. Accordingly, such controversial division in this literature strongly motivates rigorous empirical evaluations to assess the value-added of the educational credentials of teachers in learning terms.

To address these aforementioned research contentions, this study elects to focus on the case of China, where swift teacher education reforms and reorganization policies have been taking place in light of public demands for a better quality of education (Zhou and Reed, 2005). A key factor linking the Chinese context to the broader international debate on the usefulness of the educational credentials of teachers as a predictor of effective teaching has been the keen policy interest and emphasis of China in improving teacher educational attainment and the quality of teacher education (Liu and Xie, 2021). Historically, the educational credential requirements of teachers in China varied considerably by instructional level. For instance, primary schools have traditionally set high school diploma, or *shizhuan*, as a basic requirement for the entry of a new teacher, whereas secondary schools often require at least a 2-year associate's degree, or *zhuanke*, to be eligible for employment (Ingersoll, 2007).

In recent years, national motivation to increase the supply of teachers who hold higher levels of educational credentials has been strong. In 1999, the State Council initiated *Action Plan to Revitalize Education in the Twenty-first Century* and widely publicized its *Decisions to Deepen Educational Reform and Improve Quality-Oriented Education in a Holistic Way*, which served as policy guidelines in generating national consensus to improve teacher education (Zhou, 2014). As mandated in the *Ten-Five Teacher Education Reform and Development Outline*, the Ministry of Education (2002) posits that all new teachers in compulsory education should hold at least a bachelor's degree, which led to a sharp increase in the percentage of new teachers who hold bachelor-level educational credentials. In addition, the Ministry of Education initiated Free Teacher Education Program (FTEP) in 2007, recruiting more than 46,000 bachelor-degree teacher candidates in its first cycle (Qian et al., 2020). As a result, the proportion of teachers who hold bachelor's degrees drastically increased within a relatively short time frame. According to the Ministry of Education (2016), the percentage of lower-secondary teachers with at least a Bachelor of Education (BEd) degree expanded quintuple times from 12.4% of the total teaching force in 1999 to 59.4% in 2009, and by 2016, approximately 4 out of 5 lower-secondary teachers have at least a BEd degree.

Despite the rapid movement to bachelorize teachers in China, its benefits realized in pupil learning terms have been unclear. For instance, a cluster of scholars increasingly scrutinize this policy movement to improve the quality of a teacher by "bachelorizing" teachers and question whether attaining a higher level of educational credential contributes to effective teaching (Jin, 2007; Zhou et al., 2011; Hu, 2015). In general, there still exists an open debate in academic circles regarding the effects on learning, resulting from the rapid movement to bachelorize teachers, which echoes the broader international discussion regarding the usefulness of the educational credentials of teachers as a predictor of effective teaching. In particular, little is known empirically about the actual benefits of promoting mass teacher education at the bachelor level (Zhou, 2014). This study, through employing an empirical analysis leveraging pupil fixed-effect modeling, evaluate the extent to which the educational credentials of teachers may serve as a predictor of effective teaching and quantify their contributions to learning.

MATERIALS AND METHODS

Subjects

The research design in this study anchors on two key components, namely, the first being the publicly available longitudinal China Education Panel Survey (CEPS) dataset, and the second is leveraging pupil fixed-effect modeling approach. This study uses CEPS baseline (2013) and follow-up (2014)

tracked pupil panel dataset, which is collected through a school-based, multistage, multistrata, and probability proportional-to-size sampling (PPS) design. The CEPS is a nationally representative study that includes a set of five independently administered questionnaires, each of which is distributed to pupils and to their parents, teachers, and principals. The combined panel dataset contains rich information on academic achievement, demographics, educational attitudes, and teacher information. Inclusion of subjects is solely based on the availability of teacher-pupil linked data at both baseline and follow-up waves of CEPS, for which 5,032 subjects in the seventh-grade cohort are included in the analytic sample.

Measures

As with many existing studies, a common concern with analyzing pupil achievement data is the inclusion as a regressor of the lagged pupil learning outcome, because the variable is likely to be measured with error and because any correlation of learning outcomes over time would make the variable endogenous in the specification (Clotfelter et al., 2007). Consequently, this study specifies the key dependent variable as the gain of a pupil in learning outcomes from one period to the other, termed as "pupil learning gains," which is calculated as the difference in test performance on the core-subject standardized tests, which is observed between baseline and follow-up waves of CEPS. Unlike previous studies, test score information is directly validated and obtained from educational administrative authorities in the CEPS study. All raw test scores are standardized within each subject to display a mean of 70 points and SD of 10 points. The key explanatory variable, i.e., the educational credentials of teachers, is operationally determined as a binary variable, indicating whether teachers hold at least a BEd degree, or otherwise. A rich set of pupil- and teacher-level variables are analyzed as measures to control for potential confounding influence. Empirical validity and reliability information on core-subject standardized test score measures is presented in the descriptive "Results" section.

Statistical Analysis

This study utilizes a pupil fixed-effect modeling approach to relate variation in educational credentials among teachers to pupil learning gains that varies by subject-teacher pairs. This empirical analytic approach is often referred to as pupil fixed-effect modeling in the psychometrics and econometrics literature (Metzler and Woessmann, 2012), and its utility rests in addressing endogeneity concerns by eliminating confounding effects of observable and unobservable pupil-level factors that remain invariant among subject-teacher pairs. In addition, the rich panel dataset on pupil and teacher information enables the inclusion of control variable vectors, such as pupil attitude and effort as well as teacher background and experience. Operationally, the empirical pupil fixed-effect model was derived by the following mathematical formula:

$$\left[Y_{ijt} - Y^{'}_{ij(t-1)}\right] = \theta \cdot T_{j(t-1)} + \phi \cdot X_{ij(t-1)} + \delta \cdot C_{j(t-1)} + \mu_i + \varepsilon_{ijt}$$

where the key dependent variable, $Y_{ijt} - Y^{'}_{ij(t-1)}$, is estimated as the difference in the standardized test performance of pupil i for teacher-subject pair j, i.e., the observed learning gains between baseline and follow-up surveys. The key explanatory variable, i.e., the educational credentials of teachers, as indicated by $T_{j(t-1)}$ is set to equal 1 if teachers hold at least a BEd degree and set to equal 0 if otherwise. To account for pupil learning input, experience, and attitude, which may vary in relation to teacher-subject pairs, the fully specified model also includes $X_{ij(t-1)}$, which is conceived as a pupil-level vector of control variables, consisting of time-lagged measures of participation in private tutoring, frequency of pupil-teacher interaction, and affinity toward the subject. In the same vein, $C_{j(t-1)}$ represents a teacher-level vector of control variables, which includes teacher background information, such as gender, teaching experience, homeroom teaching status, teacher licensure status, teacher job rank, and receipt of teaching award, which may vary across teacher-subject pairs and confound results. Critically, by including pupil fixed-effects (μ_i), the fully specified model effectively minimizes the potentially confounding influence of pupil-level, teacher-level, and school-level factors that are invariant across subject-teacher pairs.

RESULTS

Preliminary Analysis

In **Table 1**, descriptive statistic information is provided for the final analytic sample. A key feature of the pupil fixed-effect model is that there are as many rows of observation per pupil as there are teacher-subject pair categories. Consequently, the analytic sample size in this study is 15,096 since there are three possible teacher-subject pair categories, namely, Home Language, Numeracy, and Foreign Language.

In more detail, 83.5% of teachers report holding at least a BEd degree, 73.4% of teachers are female, the average teaching experience is 14.669 years (SD = 8.668), and 29.4% of teachers are responsible as homeroom teachers. It is also worth mentioning that teachers in China are subject to official teaching licensure requirements, which are legally required unless teachers are undergoing a probational induction period (Ingersoll, 2007). There are 97.7% of teachers in the analytic sample holding official teaching license. Once teachers pass their probational induction period, they may apply through their local education authority for a series of job ranks, commonly beginning from the progressive order of rankless, level 3, level 2, level 1, senior level 2, and senior level 1 (Ministry of Education, 2014). Prior studies have found teacher job rank, or *zhicheng*, to be positively correlated with length of tenure and teaching performance (Wang and and, 2016). In this study, there are 84.0% of teachers ranked above level 2. According to Chu et al. (2015), teachers must meet rigorous requirements in order to be promoted to the next rank, which includes professional development certification and classroom performance audits, which are aimed to assess the mastery of pedagogy, instructional tools, and classroom management of teachers. Additionally, a teaching award is another indicator of teaching performance and is commonly bestowed by various educational authorities, ranging from national-, provincial-, municipal-, district-, and school levels, often through the form of pedagogical skills and

TABLE 1 | Sample descriptive statistics information (n = 15,096).

	Definition and metric	Mean
Key explanatory variable		
Hold at least a BEd	Teacher's educational credential, set to equal 1 if holding at least a BEd, and set to equal 0 if otherwise.	0.835
Teacher-level variables		
Female	Teacher's sex, set to equal 1 if female, and set to equal 0 if otherwise.	0.734
Teaching experience	Teacher's length of teaching experience, self-reported in years	14.669 (8.668)
Homeroom teaching	Teacher is responsible as homeroom teacher, set to equal 1 if homeroom teacher, and set to equal 0 if otherwise.	0.294
Hold official licensure	Teacher's official teaching licensure status, set to equal 1 if hold official license, and set to equal 0 if otherwise.	0.977
Hold job rank above level 2	Teacher's job rank, set to equal 1 if ranked above level 2, and set to equal 0 if otherwise.	0.840
Hold at least municipal teaching award	Teacher's teaching award, set to equal 1 if hold at least municipal award, and set to equal 0 if otherwise.	0.353
Pupil-level variables		
Private tutoring	Pupil is enrolled in private tutoring, set to equal 1 if enrolled in private tutoring, and set to equal 0 if otherwise.	0.146
Frequent interaction	Pupil reported frequent interaction with teacher, set to equal 1 if reported frequent interaction, and set to equal 0 if otherwise.	0.643
Subject affinity	Pupil indicated favorable attitude to subject, set to equal 1 if favorable attitude, and set to equal 0 if otherwise.	0.896

SDs in parentheses where appropriate.

classroom management competitions (Liu, 2019). In this study sample, 35.3% of teachers hold at least the municipal-level teaching award.

At the pupil level, 14.6% in the sample report as enrolled in private tutoring for Home Language, Numeracy, or Foreign Language, whereas 64.3% feel that their subject teacher frequently interacts with them, and 89.6% indicate that they hold a favorable attitude toward the subject they are learning. Finally, the distribution of the key dependent variable is plotted according to the categories of the key explanatory variable in **Figure 1**. Visually, it is observed that there is a considerable degree of first-order difference between teachers who hold at least a BEd degree and those who do not, which is depicted in terms of pupil learning gains.

In **Table 2**, correlation and descriptive statistic information on key dependent variables are presented. Correlational findings confirm that there is good first-order variation among teacher-subject pairs. The correlation coefficients of intersubject test performance at baseline and follow-up, which serve as indicators for reliability among core-subject standardized tests, are relatively high, suggesting that the three tests exhibit reasonable reliability. Specifically, the intersubject correlation coefficients at baseline are 0.641 ($p < 0.05$), 0.720 ($p < 0.05$), 0.70 ($p < 0.05$), for Home Language and Numeracy, Home Language and Foreign Language, as well as Numeracy and Foreign Language, respectively. In addition, the intersubject correlation coefficients at follow-up are 0.69 ($p < 0.05$), 0.68 ($p < 0.05$), 0.73 ($p < 0.05$), respectively, for the same list of subject pairs. For each subject, the intrasubject correlation coefficients between baseline and follow-up, which demonstrate test-retest reliability, are 0.73 ($p < 0.05$), 0.71 ($p < 0.05$), 0.74 ($p < 0.05$), for Home Language, Numeracy, and Foreign Language, respectively.

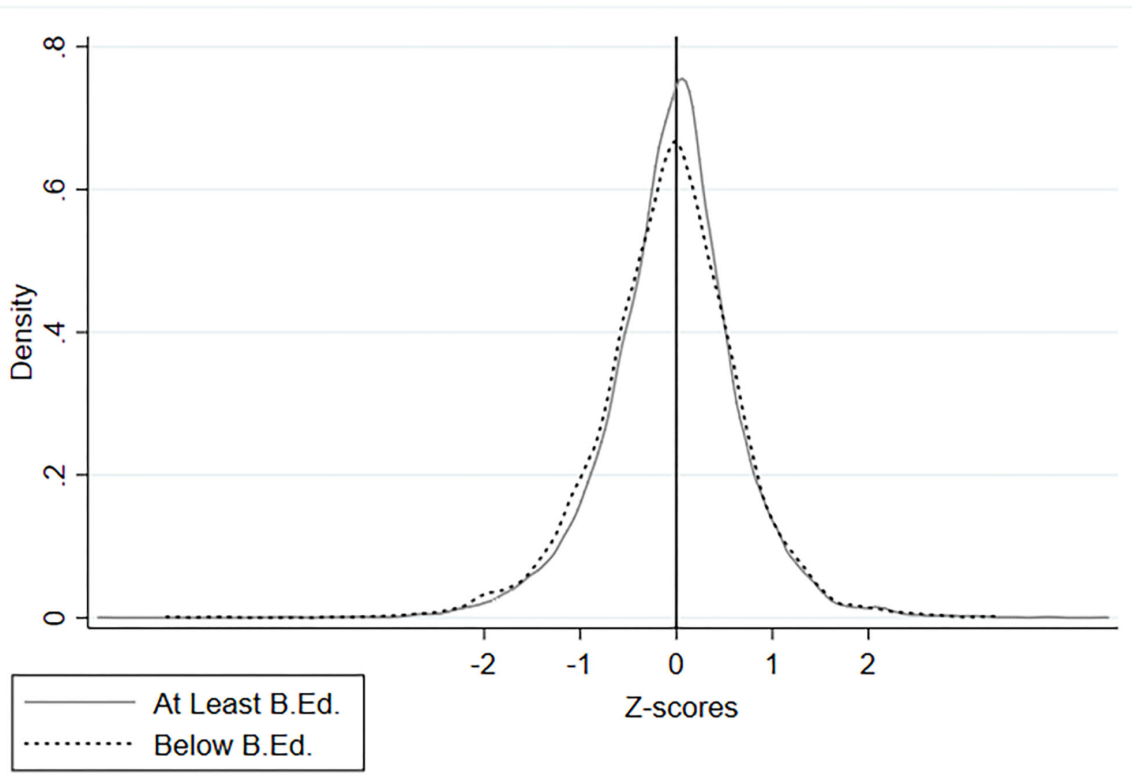


FIGURE 1 | Distribution of pupil learning gains by the educational credentials of teachers. This chart plots the difference in test performance on core-subject standardized tests between baseline and follow-up, by the educational credential of the subject teacher. Kernel density is computed using the Epanechnikov method.

TABLE 2 | Dependent variable correlation matrix and descriptive statistics.

Dependent variables	1	2	3	4	5	6		
Baselinetest	1	Home language	1					
	2	Numeracy	0.641*	1				
	3	Foreign language	0.720*	0.700*	1			
Follow-uptest	4	Home language	0.740*	0.561*	0.631*	1		
	5	Numeracy	0.583*	0.706*	0.612*	0.685*	1	
	6	Foreign language	0.584*	0.615*	0.739*	0.677*	0.732*	1
N	5032	5032	5032	5032	5032	5032		
Mean	75.64	73.73	79.86	78.59	71.53	67.57		
SD	18.22	28.07	26.62	22.06	32.76	30.19		

*Denotes $p < 0.05$.

Pupil Fixed-Effect Analysis

The main results of the fully specified pupil fixed-effect regression are shown in **Table 3**, for which the analysis focuses on evaluating the extent to which holding at least a BEd degree serves as a predictor of pupil learning gains between baseline and follow-up. Importantly, the pupil fixed-effect modeling approach limits this analysis within the same pupil, so all confounding influences of pupil-level, teacher-level, and school-level factors that are invariant across subject-teacher pairs are minimized. In more specific terms, the pupil fixed-effect regression coefficient for the key explanatory variable in this study, i.e., "Hold at least a BEd degree," is evaluated at 0.042 z-score units ($p = 0.040$), with a 95% CI between 0.008 and 0.083 z-score units.

For interpretation, this finding indicates that pupils who study with a teacher, who holds at least a BEd, are observed on average to exhibit 0.042 more SDs of learning gain between baseline and follow-up than with a teacher who holds less than a BEd degree, holding all else equal, and this result is statistically

TABLE 3 | Pupil fixed-effect regression results (n = 15,096).

Dependent variable: **Pupil learning gains (*z*-scores)**	**Coefficient**	**SE**	**95% CI**	***p*-value**
Key explanatory variable				
Hold at least a BEd degree	0.042*	0.021	[0.008, 0.083]	0.040
Pupil-level control variables				
Private tutoring	0.012	0.025	[-0.037, 0.061]	0.639
Frequent interaction	0.015	0.021	[0.026, 0.056]	0.466
Subject affinity	0.082*	0.028	[0.027, 0.138]	0.004
Teacher-level control variables				
Female	−0.022	0.017	[-0.055, 0.011]	0.184
Teaching experience	0.002	0.001	[-0.001, 0.004]	0.098
Homeroom teaching	0.058*	0.012	[0.034, 0.082]	0.001
Hold official licensure	0.008	0.045	[-0.080, 0.097]	0.856
Hold job rank above level 2	0.065*	0.030	[0.007, 0.123]	0.027
Hold at least municipal teaching award	0.001	0.016	[-0.034, 0.031]	0.935

*The adjusted R^2 of the model is 0.166. *Denotes $p < 0.05$.*

robust at the 0.05 level. To situate the size of this observed learning gain in a broader context, Evans and Yuan (2019) estimated that pupils are on average predicted to gain between 0.15 and 0.21 SDs of learning per full academic year in developing countries, while the guidance of Organisation for Economic Co-operation Development. (2016) on the number of expected learning gains per full academic year in developed countries is between 0.25 and 0.30 SDs. Conservatively speaking, findings in this analysis suggests that the number of learning that is added by a teacher who holds at least a BEd degree is approximately equivalent to at least 1 month of additional learning per full academic year.

In addition, while the pupil fixed-effect modeling approach accounts for confounding influences that may be invariant across teacher-subject pairs, there may still remain subject-varying pupil-level confounders, such as pupil attitude and effort, as well as subject-varying teacher-level confounders, such as teacher background and experience. Consequently, the full specification of the pupil fixed-effect model includes additional pupil- and teacher-level control variables. As such, the regression results for these important control variable coefficients are also worth mentioning. At the pupil level, participation in private tutoring ($p = 0.639$) and self-reported frequency of interaction ($p = 0.466$) are not statistically correlated with learning gains, whereas self-reported affinity toward the subject ($p = 0.004$) is positively associated with larger learning gains. This result suggests that the affinity of the pupil toward the subject strongly predicts better learning outcomes, even after accounting for subject-invariant factors. At the teacher level, female ($p = 0.184$), teaching experience ($p = 0.098$), official licensure ($p = 0.856$), and municipal teaching award ($p = 0.935$) are not statistically associated with the dependent variable; however, homeroom teaching ($p = 0.001$) and job rank above level 2 ($p = 0.027$) both positively predict larger learning gains. For interpretation, homeroom teaching and job rank are the two important teacher-level predictors for improved pupil learning gains, in addition to the educational credentials of teachers.

DISCUSSION AND CONCLUSION

Prior studies have indicated that there exists an open debate as to whether observable teacher background traits, such as the educational credentials of teachers, are predictive of teaching effectiveness (Chingos and Peterson, 2011). On the one hand, studies have argued that the human capital development of teachers is vital in determining how well they translate what they know into what pupils learn, particularly in making instructional experiences meaningful to pupils who may possess wide-ranging learning needs (Darling-Hammond et al., 2005). On the other hand, critics of teacher education have asserted that pre-service teacher education is only useful to the extent that teachers are minimally qualified to be a teacher but may not necessarily translate into effective teaching (Ballou and Podgursky, 2005); therefore, many ponder the policy implications since there is some indication that the higher levels of the educational credentials of teachers do not predict elevated learning gains (Hanushek and Rivkin, 2006). Not to mention that academic skeptics have scrutinized recent policy engagements in China, which promote mass teacher education at the bachelor level (Zhou, 2014; Hu, 2015).

The research objective in this study was to evaluate the extent to which the educational credentials of teachers may serve as a predictor of effective teaching and quantify their contributions to pupil's educational success. In broad strokes, this study is centered on the Chinese educational context and employs a pupil fixed-effect modeling approach to relate differences in the educational credentials of teachers to variations in pupil learning gains all within the same pupil, while eliminating problematic confounders at the pupil-, teacher-, and school levels. In broad terms, the Chinese context provides a useful analytic case where policy engagement with improving teacher educational attainment has been swift and strong. More specifically, analytic findings that are generated using baseline and follow-up CEPS datasets indicate that returns to studying under teachers holding higher levels of educational credential are substantial, as calculated in terms of pupil learning gain.

Particularly, pupils who study with teachers who hold at least a BEd degree are observed to gain 0.042 higher SDs of measured learning between baseline and follow-up, compared with when the pupil studies with teachers who do not hold such credentials. This effect translates to as much as 1–2 months of additional learning gain per academic year (Organisation for Economic Co-operation Development., 2016). Considering the potentially compounding effect over time, the number of additional learning gains that better-educated teachers can deliver could be quite significant, particularly if pupils are systematically exposed to better-educated teachers in the course of their learning careers.

Findings in this study add rigorous evidence to a growing body of literature highlighting the criticality of the human capital attainment of teachers through programs of tertiary teacher education preparation for boosting instructional effectiveness and bolstering pupil learning and development (Clotfelter et al., 2007; Winters et al., 2012). Conceptually, focusing on improving teacher education and bettering teacher preparation at the tertiary level, well before teachers arrive in the classroom, is an attractive policy option because existing evidence supports that instructional effectiveness matters for learning outcomes (Allen et al., 2013). At the theoretical level, it is crucial to recognize that the human capital development of teachers is a deterministic factor affecting that of their pupils and that how well teachers are educated reflects how well pupils can be reasonably expected to learn. The importance of the level of human capital development of teachers in meeting the increasingly complex demands of instructional success in the classroom should not be underemphasized. As such, a joint policy emphasis should be placed on improving the quality of teacher education and incentivizing teacher candidates in pursuing higher levels of educational credentials. In practice, however, many countries find it a fiscal and logistical challenge to procure large populations of highly qualified teachers, given that the personnel-related costs are already taking up significant portions of educational budgets (Baumol, 1993; Levin, 2011). To that end, while this study does not directly offer fiscal solutions, its findings critically counter against opponents of teacher education and show that better-educated teachers indeed lead to improved learning gains, which can translate into large growth prospects and societal gains in the long run (Hanushek, 2011).

Additionally, closer attention must be paid to understand the inner workings of the Chinese model of teacher education, when contextualizing and translating findings in this study to a culturally and educationally different circumstance. Most importantly, neither all teacher education programs are of the same quality nor do they offer teacher candidates the same professional preparation to make a meaningful difference in the classroom. For instance, Li (2012) observed that the Chinese model of teacher education has been historically a hybrid system that involves a combination of specialized teacher training colleges and comprehensive research universities, and the model emphasizes core values such as independence, adaptability, and diversity stemming from Confucian epistemology and pragmatism. In addition, studies elsewhere have highlighted the importance of cultural relevance in supporting teacher candidates to self-reflect, critique, and link their unique teacher education experience to effective instructional practices in the classroom (Acquah et al., 2020). Finally, in order for teacher candidates to truly benefit from high-quality teacher education programs, studies have shown that they must invest in themselves through hard work, as indicated through undergraduate course grades and course hours (Kukla-Acevedo, 2009).

Nonetheless, two important limitations in this study are also worth mentioning and can perhaps motivate future research in addressing them. First, it is critical to acknowledge that this study only evaluates the extent to which the educational credentials of teachers may serve as an indication for effective teaching, in that its contribution to pupil learning gains is empirically calculated. There are many additional dimensions of human capital development traits of teachers, which are worth considering, such as teacher certification (Harris and Sass, 2009) and teaching experience (Kane et al., 2008), as well as teacher rank in the Chinese context, where the findings in this study, as well as that in the studies by Hannum and Park (2007) and Chu et al. (2015), have shown to positively influence pupil learning. These positive signals of teacher quality are not only indicative of the human capital development of teachers but also more importantly predict their instructional effectiveness when they enter the classroom. Second, the CEPS study does not include in its research design to randomly assign teachers to pupils, and therefore, there are reasonable concerns that sample selection bias at both the school- and classroom level may be present and confound findings. For instance, if more qualified teachers are systematically assigned to teach classes enrolled with high performers, or if more qualified teachers are compensatorily assigned to teach in low-performing classes, the relationship between the educational credentials of teachers and the pupil learning gains becomes ambiguous. Notwithstanding, the pupil fixed-effect modeling approach attempted to address some of these issues to the possible extent, by leveraging within-pupil, between-subject variation, in addition to including a rich set of pupil- and teacher-level control variables. In this regard, however, the estimated effect of the educational credentials of teachers could reflect a combined effect, to some extent, comprised of unobserved peer characteristics at both pupil- and teacher levels. Consequently, future research could readily take advantage of the ongoing longitudinal CEPS and build on findings in this study to explore how sample selection bias may be at play.

DATA AVAILABILITY STATEMENT

China Education Panel Survey dataset is publicly-available at https://doi.org/10.18170/DVN/KURJUU.

ETHICS STATEMENT

Written informed consent to participate in the study was obtained by National Survey Research Center of Renmin University, PR China. All research ethics approvals were granted by the Institutional Review Board of Renmin University, PR China.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This study was financially supported by the National Social Science Foundation of China (Grant No. CJA200256).

REFERENCES

Acquah, E., Szelei, N., and Katz, H. (2020). Using modelling to make culturally responsive pedagogy explicit in preservice teacher education in Finland. *Br. Educ. Res. J.* 46, 122–139. doi: 10.1002/berj.3571

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., and Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: predicting student achievement with the Classroom Assessment Scoring System-Secondary. *School Psychol. Rev.* 42, 76–98. doi: 10.1080/02796015.2013.12087492

Altinok, N., and Kingdon, G. (2012). New evidence on class size effects: a pupil fixed effects approach. *Oxford Bull. Econ. Stat.* 74, 203–234. doi: 10.1111/j.1468-0084.2011.00648.x

Ballou, D., and Podgursky, M. (2005). Reforming teacher preparation and licensing: what is the evidence? *Teachers Coll. Rec.* 102, 5–27. doi: 10.1111/0161-4681.00046

Bastian, K. (2019). A Degree Above? The value-added estimates and evaluation ratings of teachers with a graduate degree. *Educ. Finan. Policy* 14, 652–678. doi: 10.1162/edfp_a_00261

Bau, N., and Das, J. (2017). *The Misallocation of Pay and Productivity in the Public Sector: Evidence From the Labor Market for Teachers*. Washington, D.C.: The World Bank. doi: 10.1596/1813-9450-8050

Baumol, W. (1993). Health care, education and the cost disease: a looming crisis for public choice. *Public Choice* 77, 17–28. doi: 10.1007/BF01049216

Chetty, R., Friedman, J., and Rockoff, J. (2014). Measuring the impacts of teachers II: teacher value-added and student outcomes in adulthood. *Am. Econ. Rev.* 104, 2633–2679. doi: 10.1257/aer.104.9.2633

Chingos, M., and Peterson, P. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Econ. Educ. Rev.* 30, 449–465. doi: 10.1016/j.econedurev.2010.12.010

Chu, J., Loyalka, P., Chu, J., Qu, Q., Shi, Y., Li., et al. (2015). The impact of teacher credentials on student achievement in China. *China Econ. Rev.* 36, 14–24. doi: 10.1016/j.chieco.2015.08.006

Clotfelter, C., Ladd, H., and Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Econ. Educ. Rev.* 26, 673–682. doi: 10.1016/j.econedurev.2007.10.002

Darling-Hammond, L., Holtzman, D., Gatlin, S., and Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Educ Policy Anal. Arch.* 13, 1–48. doi: 10.14507/epaa.v13n42.2005

Evans, D., and Yuan, F. (2019). Equivalent years of schooling: a metric to communicate learning gains in concrete terms. *World Bank Working Papers No. 8752*. Washington, DC: World Bank Group.

Hannum, E., and Park, A. (2007). *Education and Reform in China*. London: Routledge.

Hanushek, E. (2011). The economic value of higher teacher quality. *Econ. Educ. Rev.* 30, 466–479. doi: 10.1016/j.econedurev.2010.12.006

Hanushek, E., and Rivkin, S. (2006). Teacher quality. *Handb. Econ Educ* 2, 1051–1078. doi: 10.1016/S1574-0692(06)02018-6

Hanushek, E., and Rivkin, S. (2012). The distribution of teacher quality and implications for policy. *Ann. Rev. Econ.* 4, 131–157. doi: 10.1146/annurev-economics-080511-111001

Harris, D., and Sass, T. (2009). The effects of NBPTS-certified teachers on student achievement. *J. Policy Anal. Manage.* 28, 55–80. doi: 10.1002/pam.20402

Harris, D., and Sass, T. (2011). Teacher training, teacher quality and student achievement. *J Public Econ.* 95, 798–812. doi: 10.1016/j.jpubeco.2010.11.009

Hu, J. (2015). Review of elevation of primary school teacher's record of schooling in Chinese Mainland. *Jiangsu Educ. Res.* 262, 7–12. doi: 10.13696/j.cnki.jer1673-9094.2015.10.002

Ingersoll, R. (2007). *A Comparative Study of Teacher Preparation and Qualifications in Six Nations*. Washington, D.C.: Consortium for Policy Research in Education. doi: 10.12698/cpre.2007.rb47

Jackson, C., and Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *Am. Econ. J.* 1, 85–108. doi: 10.3386/w15202

Jin, M. (2007). Study on the Curriculum and Teaching in Teacher Education. *Stud. Fore. Educ.* 34, 19–22.

Kane, T., Rockoff, J., and Staiger, D. (2008). What does certification tell us about teacher effectiveness? Evidence from New York city. *Econ. Educ. Rev.* 27, 615–631. doi: 10.1016/j.econedurev.2007.05.005

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Econ. Educ. Rev.* 28, 49–57. doi: 10.1016/j.econedurev.2007.10.007

Levin, H. (2011). The economics of education. *Albany Govern. Law Rev.* 4, 395–425.

Li, J. (2012). The Chinese model of teacher education: retrospects and prospects over a century. *Front. Educ. China* 7, 417–442. doi: 10.1007/BF03397151

Liu, J. (2019). *The Economics of Teacher Occupational Choice in China*. Dissertation: Columbia University. doi: 10.7916/d8-mrhf-bq59

Liu, J. (2021). Exploring teacher attrition in urban china through interplay of wages and well-being. *Educ. Urban Soc.* 53, 807–830. doi: 10.1177/0013124520958410

Liu, J., and Steiner-Khamsi, G. (2020). Human capital index and the hidden penalty for non-participation in ILSAs. *Int. J. Educ. Dev.* 73, 1–9. doi: 10.1016/j.ijedudev.2019.102149

Liu, J., and Xie, J. (2021). Invisible shifts in and out of the classroom: dynamics of teacher salary and teacher supply in urban China. *Teach. Coll. Rec.* 123, 1–38. Available online at: https://www.tcrecord.org/Content.asp?ContentId=23554

Metzler, J., and Woessmann, L. (2012). The impact of teacher subject knowledge on student achievement: evidence from within-teacher within-student variation. *J. Dev. Econ.* 99, 486–496. doi: 10.1016/j.jdeveco.2012.06.002

Ministry of Education. (2002). *Ten-Five Teacher Education Reform and Development Outline*. Beijing: China Ministry of Education.

Ministry of Education. (2014). *Education Statistics Yearbook 2014*. Beijing: China Ministry of Education Press.

Ministry of Education. (2016). *Education Statistics Yearbook 2016*. Beijing: China Ministry of Education Press.

Noell, G., Burns, J., and Gansle, K. (2019). Linking student achievement to teacher preparation: emergent challenges in implementing value added assessment. *J. Teach. Educ.* 70, 128–138. doi: 10.1177/0022487118800708

Organisation for Economic Co-operation and Development. (2016). *Education at a Glance 2016: OECD Indicators*. Paris, France: OECD Publishing.

Qian, H., Youngs, P., Hu, S., and Prawat, X. (2020). Will China's Free Teacher Education Policy address teacher shortages in rural schools or reproduce existing inequality? *COMPARE* 50, 713–725. doi: 10.1080/03057925.2018.1559037

Wang, L., and, Lewin, K. (2016). Two *Decades of Basic Education in Rural China: Transitions and Challenges for Development*. London: Springer. doi: 10.1007/978-981-10-2120-6

Winters, M., Dixon, B., and Greene, J. (2012). Observed characteristics and teacher quality: Impacts of sample selection on a value added model. *Econ. Educ. Rev.* 31, 19–32. doi: 10.1016/j.econedurev.2011.07.014

Zhou, J. (2014). Teacher education changes in China: 1974–2014. *J. Educ. Teach.* 40, 507–523 doi: 10.1080/02607476.2014.956543

Zhou, J., Gong, A., and Tang, Y. (2011). Studying Undergraduate Teacher Preparation Programs. *Contempor. Teach. Educ.* 4, 13–17. doi: 10.16222/j.cnki.cte.2011.02.002

Zhou, J., and Reed, L. (2005). Chinese government documents on teacher education since the 1980s. *J. Educ. Teach.* 31, 201–213. doi: 10.1080/02607470500169030

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 02 December 2021
doi: 10.3389/fpsyg.2021.749837

How to Implement Game-Based Learning in a Smart Classroom? A Model Based on a Systematic Literature Review and Delphi Method

Liuxia Pan[1], Ahmed Tlili[2], Jiaping Li[1], Feng Jiang[1], Gaojun Shi[1], Huiju Yu[3] and Junfeng Yang[1]**

[1] Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, China, [2] Smart Learning Institute, Beijing Normal University, Beijing, China, [3] School of Marxism, Hangzhou Normal University, Hangzhou, China

Game-based learning (GBL) can allow learners to acquire and construct knowledge in a fun and focused learning atmosphere. A systematic literature review of 42 papers from 2010 to 2020 in this study showed that the current difficulties in implementing GBL in classrooms could be classified into the following categories: infrastructure, resources, theoretical guidance, teacher's capabilities and acceptance of GBL. In order to solve the above problems, the study constructs a technology enhanced GBL model, from the four parts of learning objective, learning process, learning evaluation, and smart classroom. In addition, this study adopted the Delphi method, inviting a total of 29 scholars, experts, teachers and school managers to explore how to implement GBL in smart classrooms. Finally, the technology enhanced GBL model was validated and the utilization approaches were provided at the conclusion part.

Keywords: games-based learning, smart classroom, teaching model, smart learning environment, education game

OPEN ACCESS

Edited by:
James Ko, The Education University of Hong Kong, Hong Kong SAR, China

Reviewed by:
Kai-Lin Yang, National Taiwan Normal University, Taiwan
Stamatios J. Papadakis, University of Crete, Greece

****Correspondence:***
Huiju Yu
yhj@hznu.edu.cn
Junfeng Yang
yjf@hznu.edu.cn

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *30 July 2021*
Accepted: *03 November 2021*
Published: *02 December 2021*

Citation:
Pan L, Tlili A, Li J, Jiang F, Shi G, Yu H and Yang J (2021) How to Implement Game-Based Learning in a Smart Classroom? A Model Based on a Systematic Literature Review and Delphi Method. Front. Psychol. 12:749837. doi: 10.3389/fpsyg.2021.749837

INTRODUCTION

Game-Based Learning (GBL) originated from the game research in the middle of the 1950s, and from the 1980s scholars started the research and practice of integrating games into instruction. With the popularization of electronic games and the transformation of education concepts, people gradually started accepting games as learning tools (Seaborn and Fels, 2015). The published papers on WoS (Web of Science) tagged by "Game-Based Learning" have demonstrated a rapid increase and interest in this field.

GBL refers to applying games or related elements, concepts, mechanisms or designs into learning (Deterding et al., 2011), which is a study mode that integrates educational games into school teaching and self-regulated learning. As a result, learners can get immersive learning experiences while mastering knowledge and skills.

GBL has been applied into classroom teaching. However, in terms of practice, there are still some problems, such as lack of integration between gaming and teaching, a poor balance between the enjoyment effect, and the education effect. Games are either too attractive but failing to reflect studying goals, or games can be too educational but failing to trigger interests among learners (Zhang and Liu, 2007). Some educational games simply provide learning content in a digitalized way, emphasizing memorizing facts (Villalta et al., 2011). Apart from that, being constrained to

the equipped devices and internet condition of the classroom, the effect and experience of games is much less satisfying (Shin and Chung, 2017; Halloran and Minaeva, 2019). Sometimes, due to the hardware conditions, applications of digital games have to be forgone. Many scholars and enterprises conducted related design and research of digital educational games, but its practical application is hard to meet the requirements of related studying activities because of the location, equipment, and internet (Xuqing, 2007; Hou et al., 2012). It is clear that learning resources, classroom environment and technical configuration play a vital role in the implementation of GBL (Dickey, 2011; Sabourin and Lester, 2013). However, lots of problems exist to carry out GBL in classrooms.

With the advance of educational technology, the research and practice of smart classroom became popular since 2012 (Yang et al., 2018), which utilized digital technology to support flexible pedagogies, including GBL. The smart classroom is a type of technology-enhanced classroom space to facilitate content presentation, class management, learning resources accessing, and instructional interaction by utilizing appropriate devices and software (Huang et al., 2012). With the development of research and practice on smart education, it is possible to carry out GBL in smart classrooms to overcome the above-mentioned problems.

In a smart classroom, with GBL, students could engage in learning by using quality game resources via digital or VR devices with broadband Internet access, hence enhance the digital GBL experience. Therefore, this study aims to promote GBL in classrooms by utilizing smart classroom. Specifically, this study answers the following two research questions:

(1) What are the problems of implementing GBL in classroom?
(2) How to implement GBL in smart classrooms?

RELATED WORK

Related Concepts of Game-Based Learning

Games can be divided into many different categories based on form and content (Amory et al., 1999; Tian et al., 2018). For conducting GBL research, the following three terms are always mentioned, namely "Serious game," "Educational game," and "Digital educational game." There is a certain connection intersection and difference between these three terms. Clarifying the meanings and relationships of these three types of games can determine the scope of the game in this study more clearly. In this study, GBL is considered as a type of educational activity based on digital educational games, which can also be understood as digital game-based learning (DGBL) (Perini et al., 2018; Chen et al., 2020).

The term "serious game" was first used by Abt to describe games designed for learning (Apt, 1970). In particular, Abt stated that serious games must have an educational purpose and not be played primarily for entertainment. Serious games (Apt, 1970) can teach players knowledge and skills, and at the same time, provide professional training and simulation. Serious games have a proven ability to facilitate the development of skills, abilities and attitudes due to their focus on problem-solving, to which players are exposed (The Gamification of Learning and Instruction, n.d.; Ritterfeld et al., 2009). The content of serious games involves personnel training, policy discussion, military, education, health, medical treatment, etc.

Educational games are games explicitly designed for education (Amory and Seagram, 2003; Ahmad et al., 2015). It includes both physical and digital games. Educational games in a narrow sense are electronic games specially developed for educational purposes (Moreno-Ger et al., 2008; Habgood and Ainsworth, 2011). Educational games in a broad sense not only involve traditional games (Vos et al., 2011) (such as origami, seven-piece puzzle, messaging game, etc.), but also include all educational software, teaching aids, toys with both the characteristics of education and fun, for example, electronic game tables developed for educational use, commercial games with educational value, and some interesting educational software, etc. Educational games should be developed by considering the objectives and functions of education.

Digital educational games (also referred sometimes as educational video games) are educational games which are digital (Law and Sun, 2012). From the perspective of participating in games, digital educational games need information technology equipment and various digital platforms to support the development of games (Lin and Lin, 2014; Aslan and Balci, 2015; Hawlitschek and Joeckel, 2017). Digital educational games also need to meet educational features, which can promote learners' understanding of the learning content. There are several types of digital educational games, including adventure and role-playing games, business games, board games, combat games, logic games and puzzles, and word games (Alessi and Trollip, 2001), and digital educational games may be designed for single player (Miller et al., 2011) and multi-players (Annetta et al., 2009).

Advantages of Game-Based Learning

GBL is often characterized as more fun, engaging, moving, and symbiotic (Brangier and Marache-Francisco, 2020; Osipovskaya and Miakotnikova, 2020; Tundjungsari, 2020). GBL allows learners to participate in authentic learning environments, providing a fun, interactive and challenging learning environment while enabling learners to experience and apply knowledge (Chen et al., 2018). GBL provides learners with a contextualized and personalized learning environment (Sykes and Dubreil, 2019) that meets the individual needs of different types of learners.

GBL is a type of educational game that improves students' attitudes and approaches to learning and allows them to appreciate the learning process itself (Yadav and Oyelere, 2020). Many studies have shown that digital game-based learning has a positive impact on learners' motivation, attitude (Tapingkae et al., 2020; Taub et al., 2020), engagement and performance (Eltahir et al., 2021). The use of game elements, such as levels, points, leaderboards and competitive environment, can not only promote students' external motivation, but also positively affect students' behavior and increase their internal motivation in subjects and concepts that are difficult for students to

understand (Kalogiannakis et al., 2021). GBL uses game elements and aesthetics to enhance students' motivation and promote learning (Zimmerling et al., 2019). Appropriate competition and challenge can motivate learners to learn. Games often have game mechanics such as competition, scoring, and ranking that motivate learners to win, gain a sense of accomplishment and satisfaction, and make learners highly motivated to learn (Jagušt et al., 2018).

GBL not only has a positive impact on student's learning, but also increases their self-efficacy (Wang and Zheng, 2020). Digital games with interesting storylines, clear objectives and tasks to be solved make teaching and learning more diverse and effective in increasing students' interest and learning efficiency (Yang and Lu, 2021).

GBL not only engages learners in learning, but also deepens their understanding of textbook content so they can solve more complex problems (Perini et al., 2018). Learners can explore games and find different solutions to problems; therefore, creative thinking and critical thinking can be developed (Nadolny et al., 2020). Learners can explore the game and find different problem solutions; thus, the creative thinking and critical thinking could be trained (Amory et al., 1999; Nadolny et al., 2020).

Theoretical Foundations for Game-Based Learning

In this research, 16 relatively high-quality research reviews in the last 5 years have been searched from the major databases in this field (Web of Science, EBSCO ERIC (Education Resources Information Center), IEEE Xplore and SpringerLink). After reviewing these papers, it is found that their main concerns can be summarized into the following four aspects: the effectiveness of GBL (Meredith, 2016; Byun and Joung, 2018; Hussein et al., 2019; Pellas et al., 2019; Tokac et al., 2019; Chen et al., 2020; Garcia et al., 2020; Karakoç et al., 2020; Stančin et al., 2020), the future trend of GBL (Giannakas et al., 2018; Gao et al., 2020), the influencing factors of GBL (Perttula et al., 2017; Shu and Liu, 2019), the theoretical foundations of GBL's effectiveness and its practical use (Carenys and Moya, 2016; Bakan and Bakan, 2018; Ab Jalil et al., 2020).

After synthesizing some literature reviews of predecessors, this research found that there is relevant theoretical support for GBL. Some studies suggest that the theories underlying GBL studies can be classified into three categories: learning theories, motivational theories, and others (Carenys and Moya, 2016). The behaviorism, cognitivism, humanism and constructivism (Amstutz, 1999; Guy, 1999; Merriam, 2001; Conole et al., 2004). Learning theories are the basis for the development of propositions in GBL. Each learning theory has its own representative principles, which provide theoretical guidance for GBL.

According to behaviorism, players need to know their goals and achieve these goals through stimuli–reaction process (Wu et al., 2012). Cognitivists consider learning not to be simply stimulation and reinforcement, but to involve thinking (Moore and Fitz, 1993). Cognitivism emphasizes the context-dependent nature of knowledge where learning is promoted through scaffolding for task completion. Humanism emphasizes that the learner-centered approach is the most important component and players can play games at their own pace and according to their mood (Kolb, 2014). Constructivism is probably the learning theory that offers propositions closest to GBL (Carenys and Moya, 2016). It states that learners must be provided with the tools that allow them to construct their own body of knowledge and that instructors should be facilitators who accompany them in this self-learning process. These statements are strongly linked to the learner-centered education model and the active learning proposed by GBL. In the part of model construction, this study refers to the input-process output model (Garris et al., 2002), the Play Curricular activity Reflection and Discussion (PCaRD) GBL pedagogical model (Denham, 2019), and the ARCS model (Keller, 1987).

Affordance of Smart Classroom

There has been a large amount of work on smart classrooms spanning over a wide range of research areas including information communication technology, machine learning, sensor networks, mobile computing, hardware (Lämsä et al., 2018). From the educational perspective, smart classrooms should integrate physical and virtual environments to provide blended environments for learners.

The physical environment of smart classrooms includes convenient learning facilities, high-speed Internet access, comfortable surroundings, flexible space layout, etc. (Paternò and Wulf, 2017). Convenient learning facilities include various types of learning terminals, display terminals, and real recording terminals, which can effectively support the presentation and sharing of learning content and learning results, and support the communication and interaction between teachers and learners. Smart classrooms have high-speed Internet access, equipped with relatively complete network communication facilities, including wired communication devices, wireless communication devices, stable and efficient server and controller. This can ensure a fluent game process and communication, allowing learners to have a good gaming experience. This can also allow multiple devices to operate stably at the same time to meet the requirements of all learners' participation. In order to provide learners with a comfortable classroom, sensing systems are installed in the classroom, which can control the temperature, light, sound and air quality (Torrente et al., 2008). The flexible spatial layout is mainly to provide learners with a more open venue for activities, rather than confining the space for teaching activities to closed conventional rooms (Brezovszky et al., 2019). Desks and chairs with humanized designs are provided so that learners can change their positions according to their needs, and form learning groups to facilitate teamwork and group learning activities. In addition, it also includes other related equipments that can meet the needs of teaching and learning activities, such as printing equipment, multimedia editing equipment, bookcases, shelves, etc.

The virtual environment of smart classrooms, based on cloud platforms, cloud servers, cloud computing, cloud storage, etc., is normally equipped with corresponding cloud diagnostic analysis systems to build a virtual learning space. From the perspective

of learning, the virtual environment of smart classrooms should provide the functions of learning context-aware, connecting learner's community, accessing learning resources, and personalizing learning pace (Denham, 2019). When environmental or user parameters are changing, classrooms with context awareness are able to determine the reactions based on certain rules or AI algorithms (Fang and Strobel, 2011; Allsop and Jessel, 2015). Social networking, e-learning spaces, internet and other technologies in a smart classroom connect learning participants and bridge the communication between teachers and learners, allowing to extend the interaction beyond classrooms (Chen et al., 2020), which promotes the construction of a learning community. Another important element of a smart classroom is the abundance of learning resources. The digital resource platform integrates a large amount of online data and materials for learners, and manages them by category to help learners obtain high-quality learning resources more conveniently (Denham et al., 2016). In addition, learners' personalized learning is also an essential element (Belova and Zowada, 2020). The management system in the smart classroom can provide services and feedback to learners, so that they can adjust and manage the learning pace as needed, which can promote their self-regulated learning.

According to the above sorting out of the characteristics of virtual environment and physical environment in smart classrooms, the functions of smart classrooms are as follows: (1) The learning content is flexible and diverse, and can be presented quickly, clearly and smoothly on multiple screens at the same time; (2) The comfortable surrounding and space layout can enhance learning engagement and optimize the learning experience; (3) Learners and teachers can access and download rich digital resources through multiple channels at any time; (4) Learning context-aware is intelligent, which can capture, identify and record learners' learning and psychological conditions, and promote personalized learning; (5) The interaction between learners and teachers, learners and learners, and human-machine would be facilitated; (6) Real-time feedback enables teachers to recognize learners' learning achievements more effectively, so as to make more reasonable classroom adjustments, and can also provide timely feedback for learners based on the results of the provided assessment; (7) Learning communities will be connected, to form learning groups or teams, and to promote collaborative learning; and (8) Learning process will be recorded, which is a good way for learners to reflect on their learning process and find out the problems in learning.

METHODOLOGY

The data in this study was collected through two methods: a comprehensive literature review and an expert survey. Specifically, as a first step, the findings about GBL problems were first collected from the literature based on a comprehensive literature review. Then, to further increase the validity of the constructed GBL model, it was reviewed and validated by experts using Delphi method. Each of the methods (literature review and Delphi) are discussed in the following sections.

Literature Review

This review followed Kitchenham and Charters' guideline for performing a systematic literature review (Keele, 2007) and was carried out through three phases: search strategy design, study selection, data extraction and data synthesis. Using literature review, this study identified some of the common teaching and learning problems in GBL and the affordance of smart classrooms for solving the problems.

Search Strategy

The search was conducted in databases that are well-known and well established in the field of education: Web of Science, EBSCO ERIC, IEEE Xplore and SpringerLink.

The search terms were constructed by Boolean logic as follows: "game-based learning" OR "gamification learning."

In a pilot search, it appeared that the search engines of different databases use different syntax for search strings. Therefore, the search terms were adjusted to accommodate different databases.

Study Selection

The selection process consisted of two stages. The first stage was a preliminary screening, focusing on the following exclusion criteria.

- Studies which are published before 2010. This was because the term smart classrooms started to emerge in 2010.
- Studies without an abstract or in forms other than a paper (such as a poster, presentation, idea paper, etc.).
- Studies that did not elaborate on the research method used or the obtained findings.
- Studies that are not peer-reviewed.
- Studies which are not written in English.

The search term ("game-based learning" OR "gamification learning") in the databases generated 1106 articles (Web of science:380; EBSCO ERIC:420; IEEE Xplore:252; Springer Link:54). The screening in previous stage excluded 562 articles and 544articles remained. After removing duplicate articles, 383 articles basically meet the requirements.

Then, each study was downloaded in the second stage selection, where several selection criteria (see **Table 1**) where used to identify the relevance of each study to the research questions. The application of inclusion and exclusion criteria eliminated 383 articles, leaving 42 eligible studies (see **Figure 1**).

TABLE 1 | Selection criteria in the second stage.

Inclusion criteria	Exclusion criteria
Research involves the background, conceptual interpretation or significance of GBL	Research that does not involve the background, conceptual interpretation or significance of GBL
Study points out the difficulties of implementing GBL in classrooms	Study that does not point out the difficulties of implementing GBL in classrooms

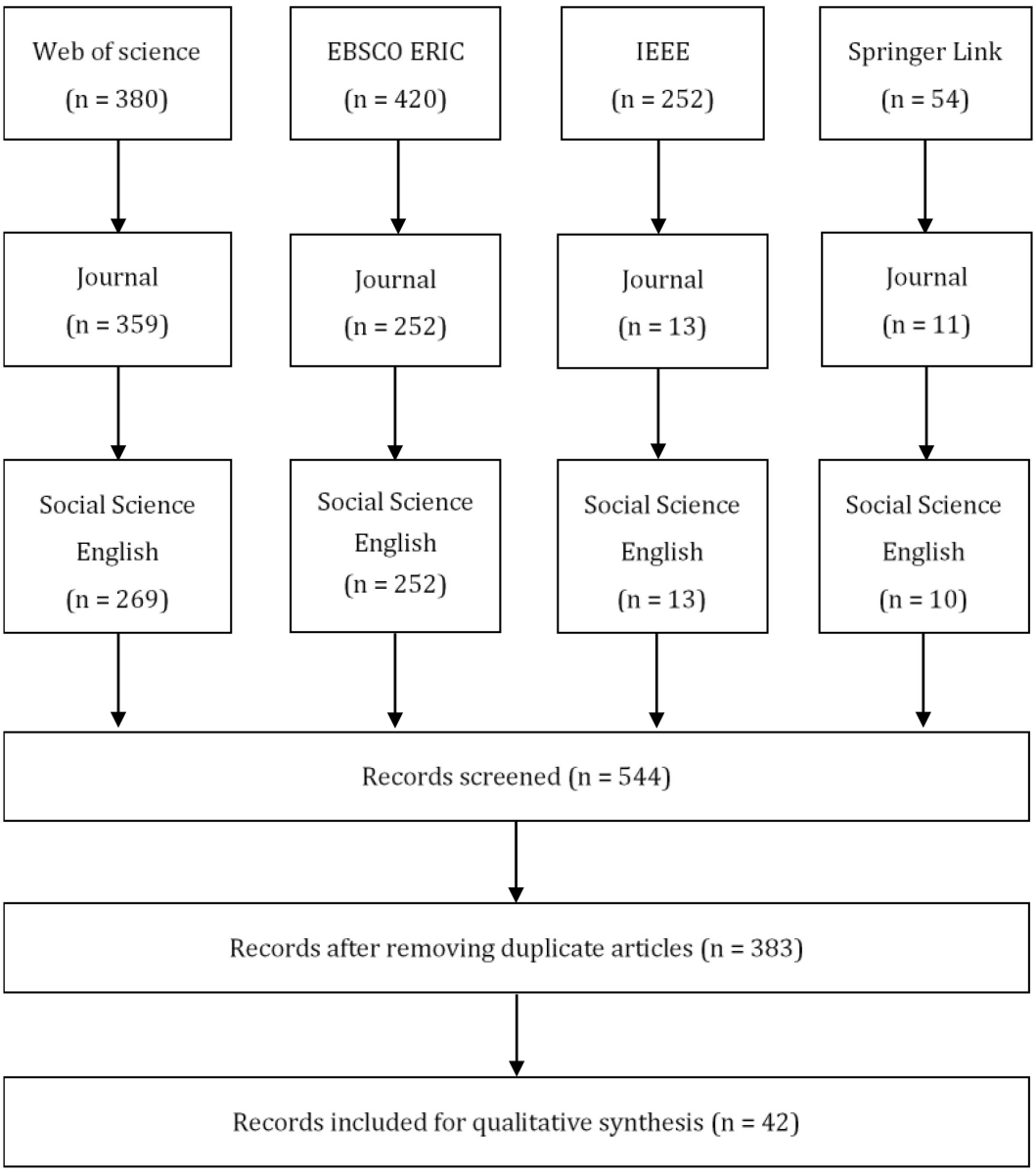


FIGURE 1 | Review process.

TABLE 2 | Coding sheet.

Database	Author(s)	Location	Title of publication	Year of publication	Type of article	Problems of implementing GBL in classrooms	Ref. (DOI/URL)

Data Extraction

An Excel form was designed to aid data extraction (**Table 2**). Each study was analyzed to derive these data, most of which were briefly presented in the results section. The analysis primarily focused on the problems of implementing GBL in classrooms.

Data Analysis

This study adopted inductive content analysis (Elo and Kyngäs, 2008) to identify the problems of implementing GBL in classrooms in the selected studies. The steps were: selecting the unit of analysis, making sense of the data and the whole, open coding, coding sheets, grouping, categorization, abstraction, and conceptual mapping.

This study arranged two researchers of this paper for the coding. Two coders performed a pilot analysis on five papers together in order to reach agreement on the semantics of "problems of implementing GBL in classrooms." Despite the inductive nature of this analysis, the coders used related literature as a reference (Lee et al., 2013; Tahir and Wang, 2020). Open

coding allowed the possibility of collecting, analyzing and categorizing other problems.

Expert Survey (Delphi Method)

A Delphi survey with GBL experts was conducted via email. Before the survey, experts were first contacted to check their interest in participating in this research. Additionally, the authors explicitly informed the experts that their participation would be anonymous. The experts were chosen based on their profiles, which should include: (1) GBL as their research interest; (2) good publication record in this area; (3) at least 5 years teaching experience.

As a result, 21 experts participated in this research (84% of active response), including scholars engaged in GBL research, and teachers who use GBL in their teaching. Despite that the experts were carefully chosen for this study to ensure the reliability of the findings, we further asked them to rate their familiarity with GBL, on a scale from 1 to 5 (where 1 is not familiar and 5 very familiar), as well as to write down their teaching experience in years. The experts had an average of 3.8 related to the familiarity with GBL, which reflect their high level of expertise and appropriateness for this study. The experts also had an average of 13 years as a teaching experience.

In the survey, the experts were requested to: (1) score 1–4 on the 25 elements extracted from the model (1 means not appropriate, 4 means very appropriate); (2) add GBL elements deemed necessary; (3) and give corresponding explanations for the choices they made. After the Delphi, we comprehensively analyzed the opinions of experts and modified the model.

RESULTS

Problems for Implementing Game-Based Learning in Classroom (Research Question 1)

Based on the conducted literature review, the following problems for applying GBL in classroom were often found. **Table 3** lists the difficulties of implementing GBL in traditional classroom mentioned in the reviewed papers. These items are classified from five aspects: infrastructure, resources, theoretical guidance, teacher's capabilities and acceptance of GBL. The classification in **Table 3** mainly relies on induction, but at the same time, it also refers to some related theoretical literature (Lee et al., 2013; Tahir and Wang, 2020).

It should be noted that the total number of papers in **Table 3** is more than the number of papers obtained by the final screening mentioned in the research method. This is because some papers have pointed out more than one type of problem, so they will be counted twice (or more) in **Table 3**. To summarize, the following problems were identified when using GBL in traditional classrooms.

(1) Digital educational games are more and more diversified, and the technologies used are more and more advanced. If teachers want to use these games to carry out GBL, they need to equip the corresponding technology and tools. However, many studies have pointed out that some of the present classroom hardware infrastructure could not support the needs of GBL, as some games with three-dimensional graphics interface have higher requirements on the central processor, memory and display card of the calculator (Nanayakkara and Whiddett, 2005; Webb et al., 2015). Traditional classrooms may be difficult to meet the needs of GBL activities. With the emergence, development and maturity of various intelligent technologies such as artificial intelligence, big data analysis, sensing technology, communication technology, cloud computing and the Internet of Things, GBL is increasingly used. The teaching practice of integrating new technologies requires a more complete learning space based on hardware facilities. It is important to establishing the infrastructure to enable gaming session (Marklund and Taylor, 2016). Therefore, one of the foundations of GBL is to have a good teaching environment, which requires appropriate technical environment to provide corresponding support.

(2) The lack of GBL resources is another major problem. The quantity and quality of educational game products need to be further improved (Sun et al., 2008). GBL needs to be based on GBL resources, such as high-quality digital games and related GBL products. However, when teachers adopt the GBL pedagogy, it is difficult for them to find the quality digital educational games. Some related enterprises and universities have begun to pay more attention to the production and development of GBL resources, and gradually strengthen the production, teaching and research integration of educational game resources development projects (Larsen, 2018; Lämsä et al., 2018; Gerodetti and Nixon, 2019; Romero et al., 2019). It will be a key research direction that can strengthen the construction of digital educational game resources, lower the threshold of GBL, and provide schools and teachers with richer products and more diversified choices. In addition to good games, the development tools for games are also what teachers need. But for now, there is still a lack of instructor-oriented authoring tools for educational games (Torrente et al., 2008; Paternò and Wulf, 2017; Brezovszky et al., 2019). It is therefore difficult for teachers to independently develop games suitable for teaching to implement GBL.

(3) There are still relatively few direct guiding theories that have a high degree of relevance for GBL. And there are few pedagogical models available for teachers who are interested in GBL (Denham, 2019). This is a major difficulty for teachers to implement GBL in classrooms. Without the guidance of proper theoretical framework, teachers may feel confused about how to apply games, what teaching activities to apply games in, how to arrange game time and learning scaffolds, how to integrate games into teaching and so on (Fang and Strobel, 2011). Not having a clear framework on GBL within the curriculum to guide teachers in the classroom, lack of subject knowledge and not knowing how to adopt new pedagogical approaches made it difficult for teachers to use games in teaching, and it also impacted on their view of teaching with games (Allsop and Jessel, 2015). Many studies have shown that it is very necessary for teachers to give them relevant theoretical guidance and instructional support (Denham et al., 2016; Belova and Zowada, 2020; Chen et al., 2020).

TABLE 3 | The difficulties of implementing GBL in classrooms.

Category	Description	Number of papers	Ref. (DOI/URL)
Infrastructure	The classroom hardware and software infrastructure is backward	4	10.1007/s10956-015-9571-7 10.3991/ijet.v14i16.10701 10.4018/ijgbl.2015010104 10.1016/j.chb.2011.11.007
	The constraints of inadequate and inappropriate technologies	5	10.1007/s11423-017-9552-z 10.1007/s10956-015-9571-7 10.1016/j.compedu.2019.04.016 10.1080/1369118X.2013.808365 10.1080/13603116.2014.885592
Resources	The quantity and quality of educational games need to be further improved	10	10.1007/s11423-017-9552-z 10.1016/j.chb.2020.106432 https://www.jstor.org/stable/jeductechsoci.17.1.42 10.1111/bjet.12346 10.1080/09639284.2016.1241951 10.3390/su12208487 10.1080/1369118X.2013.808365 10.1007/s40692-014-0008-8 10.1007/s40692-020-00174-5 10.1007/s10956-013-9436-x
	Lack of instructor-oriented authoring tools for educational games	3	10.1016/j.compedu.2018.09.011 10.1109/TLT.2011.1 10.1007/978-3-319-60291-2_14
Theoretical guidance	Lack of suitable frameworks on GBL within the curriculum	5	10.1111/bjet.12582 10.4018/ijgbl.2015010101 10.1080/09523987.2011.632277 https://www.researchgate.net/publication/343228250 10.4018/ijgbl.2015010104
	More appropriate instructional support needs to be designed to integrate games and teaching	5	10.3991/ijet.v8i6.2918 10.1007/s40299-019-00486-w 10.1007/s11528-015-0019-y 10.3390/educsci10090221 10.1109/TLT.2013.2294806
Teacher's capabilities	Teachers' instructional design ability needs to be improved	5	https://www.jstor.org/stable/26458512 10.1111/jcal.12438 10.1016/j.chb.2019.05.020 https://www.researchgate.net/publication/343228250 10.4018/ijgbl.2015010104
	Teachers' technical literacy and organizational capabilities need to be improved	3	10.3991/ijet.v9i3.3294 10.1007/s10956-015-9571-7 10.4018/ijgbl.2015010104
	Teachers need to increase the knowledge of GBL	3	10.1.1.593.1566 10.1080/09585176.2015.1018915 10.4018/ijgbl.2015010104
Acceptance of GBL	Teachers' acceptance of GBL	3	10.1016/j.compedu.2013.02.010 10.1016/j.compedu.2017.03.008 10.1007/s00530-009-0174-0
	Learners' acceptance of GBL	2	10.1177/0735633119887187 10.1111/bjet.12314
	Parents' acceptance of GBL	2	10.1016/j.compedu.2010.12.012 10.4018/ijgbl.2015010104

(4) Teachers' information literacy and GBL design capabilities need to be improved (Becker, 2007). GBL should use some software and digital games, and therefore teachers need to enhance their information literacy so that they can be able to create digital learning environments. In GBL, there are often practical problems such as insufficient integration of games and learning content, game activities deviating from learning goals, low learner participation and so on. A survey conducted in 2013, where 488 teachers were asked questions to figure out what barriers hindered them from using games in the classroom, showed that 33% of the teachers found it was difficult to integrate games into the instruction (Fishman et al., 2014). Teachers should have good information literacy to successfully blend games with instruction, and they should also have good background about educational games to solve the technical problems that may arise in the process of teaching, and to provide timely and reasonable

guidance for learners. In summary, teachers should strengthen the integration of GBL and classroom teaching, which means they need to do better in optimizing instructional design, developing diversified teaching evaluation methods, supporting learners' individualized learning and creating teaching situations (Becker, 2007; Denham et al., 2016).

(5) The acceptance of GBL is another realistic issue in the implementation of GBL. The adoption and the effectiveness of GBL depend largely on the acceptance by classroom teachers, as they can be considered the true change agents of schools. Research surveys have shown that teachers' perceptions of video games are complex. On the one hand, teachers are not really convinced that video games are very useful for enhancing their job performance. On the other hand, they believe that video games provide opportunities for learning (Bourgonjon et al., 2013; Huizenga et al., 2017). From the perspective of students, they may have a relaxed and entertaining attitude toward playing games, while ignoring the purpose of learning (Israel-Fishelson and Hershkovitz, 2020). Their level of interest in the game and the duration of the operation are also different, which will affect the participation of students (Ke et al., 2016). From the perspective of parents, they are more concerned about whether children can form a better balance between play and study life (Bourgonjon et al., 2011; Vate-U-Lan, 2015). However, what we can expect is that with the development of GBL, people's acceptance of GBL will gradually increase, and more relevant groups will have a positive view of it.

The Technology Enhanced Game-Based Learning Model (Research Question 2)

In the information age, emerging technologies could be used to help teachers implement GBL better. With the advance of educational technology, the research and practice of smart classroom became popular (Yang et al., 2018), to facilitate content presentation, class management, learning resources accessing, and instructional interaction by utilizing appropriate devices and software (Huang et al., 2012). Some studies point out the characteristics of smart classrooms include both virtual and physical environments (Rogers, 2002), provide access to data to facilitate learners' investigations (Clark et al., 2007), and produce relevant feedback for learners (Balacheff et al., 2009). By summarizing and sorting out relevant literature, this study extracts eight elements of smart classrooms:

In order to solve some of the problems (such as: infrastructure, resources, and theoretical guidance) in GBL by making good use of technology, and combined with relevant literature, this study constructs the technology enhanced GBL model. The design of GBL process in this model mainly refers to the input-process output model (Garris et al., 2002). In addition, some ideas from The Play Curricular Activity Reflection and Discussion (PCaRD) GBL pedagogical model (Denham, 2019), and the ARCS model (Keller, 1987) are also used for reference. In order to verify the validity of the model, this study adopted the Delphi method. There are 25 key elements that can be extracted from the model. All the elements of the model were identified based on a comprehensive literature review. The results of the degree of acceptance of each element of the model based on the experts' rating are shown in **Table 4**.

There are a total of 21 samples in round 1, 25 items in each questionnaire. The questionnaire's Cronbach's alpha is 0.916, and the Cronbach's alpha of each item in the questionnaire is greater than 0.9. This means that the questionnaire's reliability is high, and the collected data are reliable. The Mean score represents the expert's recognition of the elements. In this study, items below 3 are deleted. A score of 3 or below indicates that experts did not have high acceptance level toward the given element, so "22 Assigning homework" was deleted. Coefficient of Variation indicates the degree of coordination of expert degree of acceptance of the elements, the smaller the coefficient, the higher the degree of coordination of experts. It is generally believed that CV < 0.25 is a good indicator. And this study will delete items with CV ≥ 0.25, so "7 Comfortable Environment" and "25 Next Round Planning" were deleted. In addition, combined with the qualitative evaluation of experts, some elements are modified. And the final technology enhanced GBL model in this study is shown in **Figure 2**.

The technology enhanced GBL model is mainly composed of four parts: smart classrooms, GBL objectives, GBL process, and GBL evaluation. The learning objectives of a class need to be achieved through the dynamic interaction of teaching/learning and evaluation. Teachers should prepare for the GBL process by

TABLE 4 | Results of the degree of acceptance about elements of the model (the technology enhanced GBL model) based on the experts' rating.

	Element	Mean	*SD*	Coefficient of variation
1	Connecting learner's community	3.48	0.81	0.23
2	Intelligence test and data acquisition	3.57	0.68	0.19
3	Real-time feedback	3.71	0.64	0.17
4	Personalizing learning pace	3.57	0.68	0.19
5	Convenient learning facilities	3.52	0.60	0.17
6	High speed Internet access	3.57	0.51	0.14
7	Comfortable surroundings	3.00	1.00	0.33
8	Flexible space layout	3.19	0.75	0.23
9	Multimodal learning analysis	3.71	0.56	0.15
10	Pre-analysis	3.33	0.80	0.24
11	Game selection	3.62	0.74	0.20
12	Context design	3.76	0.44	0.12
13	Activity design	3.95	0.22	0.06
14	Learning contents	3.43	0.75	0.22
15	Features of the Game	3.57	0.81	0.23
16	Gamification of learning contexts	3.76	0.44	0.12
17	Thinking and inspiration	3.43	0.81	0.24
18	Gamification exploration	3.48	0.87	0.24
19	Collaboration and communication	3.62	0.50	0.14
20	Presentation and sharing	3.38	0.67	0.20
21	Learning outcomes	3.38	0.74	0.22
22	Assigning homework	2.95	0.67	0.23
23	Personalized guidance	3.71	0.46	0.12
24	Reflection and improvement	3.43	0.75	0.22
25	Next round planning	3.19	0.81	0.26

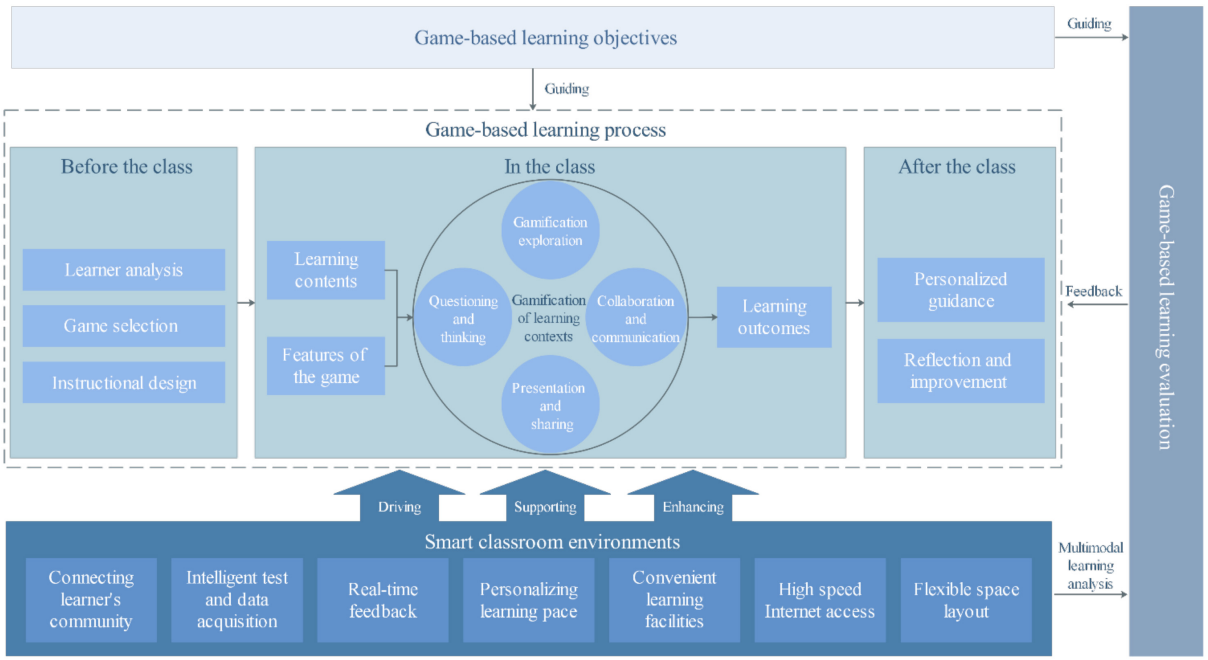


FIGURE 2 | The technology enhanced GBL model.

considering the real-time feedback to design the context, choose games, and guide learning activities. Using different technologies in the learning environment can accessing digital game resources, timely test and feedback, displaying learning analytic infographic, etc. for driving teacher's instructional design, supporting learning activities, and enhancing communications between teachers and students.

DISCUSSION, CONCLUSION, AND FUTURE RESEARCH

This study focused on the problems faced by GBL in the implementation process, and attempts to find ways to deal with some of these problems from the perspective of using technology. The study found that there were five common problems in the implementation of GBL in the classroom: (1) the backward classroom infrastructure with inadequate and inappropriate technologies, (2) the lack of quality educational game resources and instructor-oriented authoring tools, (3) the weak theoretical guidance of frameworks, curriculum and instructional support, (4) the incompetence of teacher's information literacy for GBL, (5) the stakeholder's hesitation in adopting GBL.

Based on the experts' inputs using the Delphi method, the eight elements of connecting learner's community, intelligence test and data acquisition, real-time feedback, personalizing learning pace, convenient learning facilities, high speed internet access, comfortable surroundings and flexible space layout of smart classrooms were identified (Kariippanon et al., 2020; Midcalf and Boatwright, 2020; Wang et al., 2021, p. 19). Combined with the elements and the general process of GBL, the technology enhanced GBL model was constructed. This model consisted of four parts: GBL objectives, GBL process, GBL evaluation and smart classrooms. The model explains the process and main activities of GBL from the three stages of before the class (Becker, 2007; Huang et al., 2019), in the class (Garris et al., 2002; Uzelac et al., 2015; Denham et al., 2016; Paudel et al., 2020; Kim et al., 2021) and after the class (Bayirtepe and Tuzun, 2007; Suo et al., 2008; Yang and Huang, 2015; Aguilar et al., 2020).

The design and formulation of the model can also respond to the lack of theoretical guidance to a certain extent.

(1) For the problem of infrastructure, this model provides a method for constructing suitable environments for GBL. The environments should have high-speed Internet access, which makes the game process and communication smooth. Convenient learning facilities include various types of learning terminals, display terminals, and real recording terminals, which can effectively support the presentation and sharing of learning content and learning results, and support the communication and interaction between teachers and learners. The flexible spatial layout is mainly to provide learners with a more open venue for activities, rather than confining the space for teaching activities to closed conventional rooms. It is convenient for teachers to arrange the seats of students according to different game forms and teaching activities. Desks and chairs with humanized designs are provided so that learners can change their positions according to their needs, and form learning groups to facilitate teamwork and group learning activities.

(2) For the problem of theoretical guidance, this model provides guidance for teachers' to implement GBL activities. Using this model, teachers who do not know how to implement GBL can first have a clear cognition of GBL, and can

understand the general process of GBL. In addition, teachers with GBL experience may be able to make some new discoveries and try to optimize learning analysis, learning activities and learning evaluation by using various technologies in the learning environment. This can help them to attract learners' interest and promote learners' effective learning. Therefore, the proposed model gives teachers some guidance in theory.

However, the model could not handle the other three problems of resources, teacher's capability and acceptance to GBL, which could be targeted in future studies.

The model could be used by researchers, teachers, and school administrators, or other stakeholders. For researchers, it can serve as a reference for further research on implementing GBL in smart classrooms, because in the increasingly intelligent environment, GBL will develop to a new stage, which requires researchers to carry out research to keep pace with the times. For teachers, it provides a guidance on implementing GBL in smart classrooms, because the model proposed in this paper is mainly designed according to the teaching process, teachers can refer to it in different teaching links.

(1) Before the class, teachers can choose the appropriate games and design teaching activities. Teachers can also design realistic and interactive game-based learning situations.

(2) In the class, teachers can create immersive GBL experience that can evoke thinking, promote learning by exploring through different game activities, as well as develop collaborative capability and improve interpersonal communication skills. Encouraging presentation and sharing, learners share their learning results with others and display their works through various content presentation methods in the smart classroom, such as multi-screen display and file transfer between terminals.

(3) After the class, teachers can monitor the online learning process to better enhance learning and improve the quality of teaching. Enhancing personalized guidance, where teachers can find students who have difficulty in learning by viewing and analyzing the student data collected by in the learning process. Teachers provide targeted guidance to learners to solve students' learning difficulties. Boosting reflection and improvement, where teachers reflect on the effects of teaching, redesign and improve the deficiencies. Teachers can get enlightenment from the reflection, which can become the experience and basis for teachers to improve their teaching ability.

Although the previous research basis on smart classrooms and the systematic literature review on GBL provided solid foundation for the reliability of the proposed model, this model was still in the stage of theoretical conception and had not been applied in practice. However, the idea of this model was presented at an international conference on GBL, where lots of teachers expressed that they were inspired and were willing to carry out relevant practice. Besides, the study also verified the validity of the model through the Delphi method.

This study mainly constructs a GBL model supported by smart classrooms from a theoretical perspective, however it must take further exploration in the educational field to enhance the validity. It is promising that in the near future, the integration of the GBL and smart classrooms will be explored in-depth from both theoretical and practical perspectives.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

LP and JY: conceptualization. JY, JL and FJ: methodology. HY and JY: supervision. JY, HY, JL, and GS: resources. LP, HY, and JY: investigation, data curation, writing—original draft preparation. AT, JY, HY, JL, FJ, and GS: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research work was supported by 2021 Zhejiang Provincial Philosophy and Social Planning Project (No: 21GXSZ030YB).

REFERENCES

Ab Jalil, H., Nasharuddin, N. A., Marlisah, E., Nazan, A. I. N. M., Ismail, I. A., Ma'rof, A. M., et al. (2020). Systematic Review of Enjoyment Element in Health-Related Game-Based Learning. *Int. J. Emerg. Technol. Learn.* 15:40. doi: 10.3991/ijet.v15i21.17345

Aguilar, J., Altamiranda, J., and Diaz, F. (2020). Specification of a Managing Agent of Emergent Serious Games for a Smart Classroom. *IEEE Lat. Am. Trans.* 18, 51–58. doi: 10.1109/TLA.2020.9049461

Ahmad, M., Rahim, L. A., and Arshad, N. I. (2015). An analysis of educational games design frameworks from software engineering perspective. *J. Inf. Commun. Technol.* 14, 123–151. doi: 10.32890/jict2015.14.8

Alessi, S. M., and Trollip, S. R. (2001). *Multimedia For Learning: methods and Development*. Boston: Allyn and Bacon.

Allsop, Y., and Jessel, J. (2015). Teachers' Experience and Reflections on Game-Based Learning in the Primary Classroom: views from England and Italy. *Int. J. Game Based Learn.* 5, 1–17. doi: 10.4018/ijgbl.2015010101

Amory, A., Naicker, K., Vincent, J., and Adams, C. (1999). The use of computer games as an educational tool: identification of appropriate game types and game elements. *Br. J. Educ. Technol.* 30, 311–321. doi: 10.1111/1467-8535.00121

Amory, A., and Seagram, R. (2003). Educational game models: conceptualization and evaluation ?: the practice of higher education. *S. Afr. J. High. Educ.* 17, 206–217. doi: 10.10520/EJC36981

Amstutz, D. D. (1999). Adult Learning: moving toward More Inclusive Theories and Practices. *New Dir. Adult Cont. Educ.* 82, 19–32. doi: 10.1002/ace.8202

Annetta, L. A., Minogue, J., Holmes, S. Y., and Cheng, M.-T. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Comput. Educ.* 53, 74–85. doi: 10.1016/j.compedu.2008.12.020

Apt, C. C. (1970). *Serious games: the art and science of games that simulate life in industry, government and education*. Viking: New York.

Aslan, S., and Balci, O. (2015). GAMED: digital educational game development methodology. *Simulation* 91, 307–319. doi: 10.1177/0037549715572673

Bakan, U., and Bakan, U. (2018). Estudios sobre aprendizaje basado en juegos en revistas educativas: una revisión sistemática de tendencias recientes. *Actualidades Pedagógicas* 1, 119–145. doi: 10.19052/ap.5245

Balacheff, N., Ludvigsen, S., de Jong, T., Lazonder, A., and Barnes, S. (eds) (2009). *Technology-Enhanced Learning*. Netherlands: Springer, doi: 10.1007/978-1-4020-9827-7

Bayirtepe, E., and Tuzun, H. (2007). The effects of game-based learning environments on students' achievement and self-efficacy in a computer course. *Eğitim Fakültesi Dergisi* 33, 41–54.

Becker, K. (2007). Digital game-based learning once removed: teaching teachers. *Br. J. Educ. Technol.* 38, 478–488. doi: 10.1111/j.1467-8535.2007.00711.x

Belova, N., and Zowada, C. (2020). Innovating Higher Education via Game-Based Learning on Misconceptions. *Educ. Sci.* 10:221. doi: 10.3390/educsci10090221

Bourgonjon, J., De Grove, F., De Smet, C., Van Looy, J., Soetaert, R., and Valcke, M. (2013). Acceptance of game-based learning by secondary school teachers. *Comput. Educ.* 67, 21–35. doi: 10.1016/j.compedu.2013.02.010

Bourgonjon, J., Valcke, M., Soetaert, R., de Wever, B., and Schellens, T. (2011). Parental acceptance of digital game-based learning. *Comput. Educ.* 57, 1434–1444. doi: 10.1016/j.compedu.2010.12.012

Brangier, E., and Marache-Francisco, C. (2020). "Measure of the Lived and Functional Effects of Gamification: an Experimental Study in a Professional Context," in *Advances in Ergonomics in Design*, eds F. Rebelo and M. M. Soares (Heidelberg: Springer International Publishing), 242–253. doi: 10.1007/978-3-030-20227-9_22

Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G., Pongsakdi, N., et al. (2019). Effects of a mathematics game-based learning environment on primary school students' adaptive number knowledge. *Comput. Educ.* 128, 63–74. doi: 10.1016/j.compedu.2018.09.011

Byun, J., and Joung, E. (2018). Digital game-based learning for K–12 mathematics education: a meta-analysis. *Sch. Sci. Math.* 118, 113–126. doi: 10.1111/ssm.12271

Carenys, J., and Moya, S. (2016). Digital game-based learning in accounting and business education. *Account. Educ.* 25, 598–651. doi: 10.1080/09639284.2016.1241951

Chen, C.-H., Huang, K., and Liu, J.-H. (2020). Inquiry-Enhanced Digital Game-Based Learning: effects on Secondary Students' Conceptual Understanding in Science, Game Performance, and Behavioral Patterns. *Asia Pac. Educ. Res.* 29, 319–330. doi: 10.1007/s40299-019-00486-w

Chen, C.-H., Liu, J.-H., and Shou, W.-C. (2018). How Competition in a Game-based Science Learning Environment Influences Students' Learning Achievement, Flow Experience, and Learning Behavioral Patterns. *J. Educ. Technol. Soc.* 21, 164–176.

Clark, D. B., Stegmann, K., Weinberger, A., Menekse, M., and Erkens, G. (2007). "Technology-Enhanced Learning Environments to Support Students' Argumentation," in *Argumentation in Science Education: perspectives from Classroom-Based Research*, eds S. Erduran and M. P. Jiménez-Aleixandre (Netherlands: Springer), 217–243. doi: 10.1007/978-1-4020-6670-2_11

Conole, G., Dyke, M., Oliver, M., and Seale, J. (2004). Mapping pedagogy and tools for effective learning design. *Comput. Educ.* 43, 17–33. doi: 10.1016/j.compedu.2003.12.018

Denham, A. R. (2019). Using the PCaRD digital game-based learning model of instruction in the middle school mathematics classroom: a case study. *Br. J. Educ. Technol.* 50, 415–427. doi: 10.1111/bjet.12582

Denham, A. R., Mayben, R., and Boman, T. (2016). Integrating game-based learning initiative: increasing the usage of game-based learning within K-12 classrooms through professional learning groups. *TechTrends* 60, 70–76. doi: 10.1007/s11528-015-0019-y

Deterding, S., Dixon, D., Khaled, R., and Nacke, L. (2011). "From game design elements to gamefulness: defining" gamification," in *MindTrek 2011*, eds A. Lugmayr, H. Franssila, C. Safran, and I. Hammouda (New York: ACM), 9–15. doi: 10.1145/2181037.2181040

Dickey, M. D. (2011). Murder on Grimm Isle: the impact of game narrative design in an educational game—based learning environment. *Br. J. Educ. Technol.* 42, 456–469. doi: 10.1111/j.1467-8535.2009.01032.x

Elo, S., and Kyngäs, H. (2008). The qualitative content analysis process. *J. Adv. Nurs.* 62, 107–115. doi: 10.1111/j.1365-2648.2007.04569.x

Eltahir, M. E., Alsalhi, N. R., Al-Qatawneh, S., AlQudah, H. A., and Jaradat, M. (2021). The impact of game-based learning (GBL) on students' motivation, engagement and academic performance on an Arabic language grammar course in higher education. *Educ. Inf. Technol.* 26, 3251–3278. doi: 10.1007/s10639-020-10396-w

Fang, J., and Strobel, J. (2011). How ID models help with game-based learning: an examination of the gentry model in a participatory design project. *Educ. Media Int.* 48, 287–306. doi: 10.1080/09523987.2011.632277

Fishman, B., Riconscente, M., Snider, R., Tsai, T., and Plass, J. (2014). *Empowering educators: supporting student progress in the classroom with digital games*. Ann Arbor, MI: University of Michigan.

Gao, F., Li, L., and Sun, Y. (2020). A systematic review of mobile game-based learning in STEM education. *Educ. Technol. Res. Dev.* 68, 1791–1827. doi: 10.1007/s11423-020-09787-0

Garcia, I., Pacheco, C., Méndez, F., and Calvo—Manzano, J. A. (2020). The effects of game—based learning in the acquisition of 'soft skills' on undergraduate software engineering courses: a systematic literature review. *Comput. Appl. Eng. Educ.* 28, 1327–1354. doi: 10.1002/cae.22304

Garris, R., Ahlers, R., and Driskell, J. E. (2002). Games, Motivation, and Learning: a Research and Practice Model. *Simul. Gaming* 33, 441–467. doi: 10.1177/1046878102238607

Gerodetti, N., and Nixon, D. (2019). "Exploration: using Play to Design Play—Gamification and Student Involvement in the Production of Games-Based Learning Resources for Research Methods Teaching," in *The Power of Play in Higher Education*, eds A. James and C. Nerantzi (Berlin: Springer), 283–290. doi: 10.1007/978-3-319-95780-7_36

Giannakas, F., Kambourakis, G., Papasalouros, A., and Gritzalis, S. (2018). A critical review of 13 years of mobile game-based learning. *Educ. Technol. Res. Dev.* 66, 341–384. doi: 10.1007/s11423-017-9552-z

Guy, T. C. (1999). *Providing culturally relevant adult education: a challenge for the twenty-first century*. San Francisco: Jossey-Bass.

Habgood, M. J., and Ainsworth, S. E. (2011). Motivating children to learn effectively: exploring the value of intrinsic integration in educational games. *J. Learn. Sci.* 20, 169–206. doi: 10.1080/10508406.2010.508029

Halloran, J., and Minaeva, A. (2019). Touch and play? Investigating the value of touchscreens for gamer experience. The Effects of Input Modality and Story-Based Knowledge on Users'. *Game Exp.* 32:100312. doi: 10.1016/j.entcom.2019.100312

Hawlitschek, A., and Joeckel, S. (2017). Increasing the effectiveness of digital educational games: the effects of a learning instruction on students' learning, motivation and cognitive load. *Comput. Hum. Behav.* 72, 79–86. doi: 10.1016/j.chb.2017.01.040

Hou, J., Nam, Y., Peng, W., and Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Comput. Hum. Behav.* 28, 617–623. doi: 10.1016/j.chb.2011.11.007

Huang, L.-S., Su, J.-Y., and Pao, T.-L. (2019). A Context Aware Smart Classroom Architecture for Smart Campuses. *Appl. Sci.* 9:1837. doi: 10.3390/app9091837

Huang, R., Hu, Y., Yang, J., and Xiao, G. (2012). The concept and characters of smart classroom. *Open Educ. Res.* 18, 22–27.

Huizenga, J. C., ten Dam, G. T. M., Voogt, J. M., and Admiraal, W. F. (2017). Teacher perceptions of the value of game-based learning in secondary education. *Comput. Educ.* 110, 105–115. doi: 10.1016/j.compedu.2017.03.008

Hussein, M. H., Ow, S. H., Cheong, L. S., Thong, M., and Ebrahim, N. A. (2019). Effects of Digital Game-Based Learning on Elementary Science Learning: a Systematic Review. *IEEE Access* 7, 62465–62478. doi: 10.1109/ACCESS.2019.2916324

Israel-Fishelson, R., and Hershkovitz, A. (2020). Persistence in a Game-Based Learning Environment: the Case of Elementary School Students Learning Computational Thinking. *J. Educ. Comput. Res.* 58, 891–918. doi: 10.1177/0735633119887187

Jagušt, T., Botièki, I., and So, H.-J. (2018). Examining competitive, collaborative and adaptive gamification in young learners' math learning. *Comput. Educ.* 125, 444–457. doi: 10.1016/j.compedu.2018.06.022

Kalogiannakis, M., Papadakis, S., and Zourmpakis, A.-I. (2021). Gamification in Science Education. A Systematic Review of the Literature. *Educ. Sci.* 11:22. doi: 10.3390/educsci11010022

Karakoç, B., Eryılmaz, K., Özpolat, E. T., and Yıldırım, I. (2020). The effect of game-based learning on student achievement: a meta-analysis study. *Technol. Knowl. Learn.* 1, 1–16. doi: 10.1007/s10758-020-09471-5

Kariippanon, K. E., Cliff, D. P., Okely, A. D., and Parrish, A. M. (2020). The 'why'and 'how'of flexible learning spaces: a complex adaptive systems analysis. *J. Educ. Change* 21, 569–593. doi: 10.1007/s10833-019-09364-0

Ke, F., Xie, K., and Xie, Y. (2016). Game−based learning engagement: a theory−and data−driven exploration. *Br. J. Educ. Technol.* 47, 1183–1201. doi: 10.1111/bjet.12314

Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Princeton, New Jersey, USA: Citeseer.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *J. Instr. Dev.* 10, 2–10. doi: 10.1007/BF02905780

Kim, B. J., Ke, F., Moon, J., and West, L. (2021). "Designing Dynamic Learning Supports for Game and Simulation-Based Learning in STEM Education," in *Game-based Learning Across the Disciplines. Advances in Game-Based Learning*, eds C. Aprea and D. Ifenthaler (Cham: Springer), doi: 10.1007/978-3-030-75142-5_8

Kolb, D. A. (2014). *Experiential learning: experience as the source of learning and development*. Upper Saddle River: FT press.

Lämsä, J., Hämäläinen, R., Aro, M., Koskimaa, R., and Äyrämö, S. (2018). Games for enhancing basic reading and maths skills: a systematic review of educational game design in supporting learning by people with learning disabilities. *Br. J. Educ. Technol.* 49, 596–607. doi: 10.1111/bjet.12639

Larsen, L. J. (2018). Juicing the game design process: towards a content centric framework for understanding and teaching game design in higher education. *Educ. Media Int.* 55, 231–254. doi: 10.1080/09523987.2018.1512445

Law, E. L.-C., and Sun, X. (2012). Evaluating user experience of adaptive digital educational games with Activity Theory. *Int. J. Hum. Comput. Stud.* 70, 478–497. doi: 10.1016/j.ijhcs.2012.01.007

Lee, Y., Chu, H., and Shen, C. (2013). A SURVEY RESEARCH ON THE USAGE AND THE INFLUENCEING FACTORS OF GAME-BASED LEARNING AMONG ELEMENTARY TEACHERS IN TAIWAN. *Res. Pract. Technol. Enhanc. Learn.* 8, 205–223.

Lin, H.-W., and Lin, Y.-L. (2014). Digital educational game value hierarchy from a learners' perspective. *Comput. Hum. Behav.* 30, 1–12. doi: 10.1016/j.chb.2013.07.034

Marklund, B. B., and Taylor, A.-S. A. (2016). Educational Games in Practice: the challenges involved in conducting a game-based curriculum. *Elect. J. Elearn.* 14, 122–135.

Meredith, T. R. (2016). Game-Based Learning in Professional Development for Practicing Educators: a Review of the Literature. *TechTrends* 60, 496–502. doi: 10.1007/s11528-016-0107-7

Merriam, S. B. (2001). Andragogy and self-directed learning: pillars of adult learning theory. *New Dir. Adult Cont. Educ.* 2001:3. doi: 10.1002/ace.3

Midcalf, L., and Boatwright, P. (2020). Teacher and Parent Perspectives of the Online Learning Environment Due to COVID-19. *Delta Kappa Gamma Bull.* 87:128.

Miller, L. M., Chang, C.-I., Wang, S., Beier, M. E., and Klisch, Y. (2011). Learning and motivational impacts of a multimedia science game. *Comput. Educ.* 57, 1425–1433. doi: 10.1016/j.compedu.2011.01.016

Moore, P., and Fitz, C. (1993). Gestalt theory and instructional design. *J. Tech. Writ. Commun.* 23, 137–157. doi: 10.2190/G748-BY68-L83T-X02J

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., and Fernández-Manjón, B. (2008). Educational game design for online education. *Comput. Hum. Behav.* 24, 2530–2540. doi: 10.1016/j.chb.2008.03.012

Nadolny, L., Valai, A., Cherrez, N. J., Elrick, D., Lovett, A., and Nowatzke, M. (2020). Examining the characteristics of game-based learning: a content analysis and design framework. *Comput. Educ.* 156:103936. doi: 10.1016/j.compedu.2020.103936

Nanayakkara, C., and Whiddett, D. (2005). "A model of user acceptance of e-learning technologies: a case study of a Polytechnic in New Zealand," in *Information systems technology and its applications, ISTA' 2005*, eds R. Kaschek, H. C. Mayr, and S. Liddle (Bonn: Gesellschaft für Informatik e.V), 180–189.

Osipovskaya, E., and Miakotnikova, S. (2020). "Using Gamification in Teaching Public Relations Students," in *The Challenges of the Digital Transformation in Education*, eds M. E. Auer and T. Tsiatsos (Heidelberg: Springer International Publishing), 685–696. doi: 10.1007/978-3-030-11932-4_64

Paternò, F., and Wulf, V. (2017). *New perspectives in end-user development*. Cham: Springer.

Paudel, P., Kim, S., Park, S., and Choi, K.-H. (2020). A Context-Aware IoT and Deep-Learning-Based Smart Classroom for Controlling Demand and Supply of Power Load. *Electronics* 9:1039. doi: 10.3390/electronics9061039

Pellas, N., Fotaris, P., Kazanidis, I., and Wells, D. (2019). Augmenting the learning experience in primary and secondary school education: a systematic review of recent trends in augmented reality game-based learning. *Virtual Real.* 23, 329–346. doi: 10.1007/s10055-018-0347-2

Perini, S., Luglietti, R., Margoudi, M., Oliveira, M., and Taisch, M. (2018). Learning and motivational effects of digital game-based learning (DGBL) for manufacturing education –The Life Cycle Assessment (LCA) game. *Comput. Ind.* 102, 40–49. doi: 10.1016/j.compind.2018.08.005

Perttula, A., Kiili, K., Lindstedt, A., and Tuomi, P. (2017). Flow experience in game based learning—A systematic literature review. *Int. J. Serious Games* 4, 57–72. doi: 10.17083/ijsg.v4i1.151

Ritterfeld, U., Cody, M., and Vorderer, P. (2009). *Serious games: mechanisms and effects*. Milton Park: Routledge. doi: 10.4324/9780203891650

Rogers, P. L. (ed.) (2002). *Designing instruction for technology-enhanced learning*. Hershey, PA: Idea Group Publishing. doi: 10.4018/978-1-93070-828-0

Romero, M., Arnab, S., De Smet, C., Mohamad, F., Minoi, J.-L., and Morini, L. (2019). Assessment of Co-Creativity in the Process of Game Design. *Elect. J. Elearn.* 17, 199–206. doi: 10.34190/JEL.17.3.003

Sabourin, J. L., and Lester, J. C. (2013). Affect and engagement in Game-Based Learning environments. *IEEE Trans. Affect. Comput.* 5, 45–56. doi: 10.1109/T-AFFC.2013.27

Seaborn, K., and Fels, D. I. (2015). Gamification in theory and action: a survey. *Int. J. Hum. Comput. Stud.* 74, 14–31. doi: 10.1016/j.ijhcs.2014.09.006

Shin, D.-H., and Chung, K. (2017). The effects of input modality and story-based knowledge on users' game experience. *Comput. Hum. Behav.* 68, 180–189. doi: 10.1016/j.chb.2016.11.030

Shu, L., and Liu, M. (2019). Student Engagement in Game-Based Learning: a Literature Review from 2008 to 2018. *J. Educ. Multimed. Hypermedia* 28, 193–215.

Stančin, K., Hoić-Božić, N., and Skočić Mihić, S. (2020). Using Digital Game-Based Learning for Students with Intellectual Disabilities—A Systematic Literature Review. *Inform. Educ.* 19, 323–341. doi: 10.15388/infedu.2020.15

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y., and Yeh, D. (2008). What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction. *Comput. Educ.* 50, 1183–1202. doi: 10.1016/j.compedu.2006.11.007

Suo, Y., Miyata, N., Morikawa, H., Ishida, T., and Shi, Y. (2008). Open smart classroom: extensible and scalable learning system in smart space using web service technology. *IEEE Trans. Knowl. Data Eng.* 21, 814–828. doi: 10.1109/TKDE.2008.117

Sykes, J. M., and Dubreil, S. (2019). "Pragmatics Learning in Digital Games and Virtual Environments," in *In The Routledge Handbook of Second Language Acquisition and Pragmatics*, ed. N. Taguchi (Milton Park: Routledge). doi: 10.4324/9781351164085-25

Tahir, R., and Wang, A. I. (2020). Codifying Game-Based Learning: development and Application of LEAGUÊ Framework for Learning Games. *Elect. J. Elearn.* 18, 69–87. doi: 10.34190/EJEL.20.18.1.006

Tapingkae, P., Panjaburee, P., Hwang, G.-J., and Srisawasdi, N. (2020). Effects of a formative assessment-based contextual gaming approach on students' digital citizenship behaviours, learning motivations, and perceptions. *Comput. Educ.* 159:103998. doi: 10.1016/j.compedu.2020.103998

Taub, M., Sawyer, R., Smith, A., Rowe, J., Azevedo, R., and Lester, J. (2020). The agency effect: the impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Comput. Educ.* 147:103781. doi: 10.1016/j.compedu.2019.103781

The Gamification of Learning and Instruction (n.d.). *The Gamification of Learning and Instruction*. Available Online at: https://sc.panda321.com/extdomains/

books.google.com/books/about/The_Gamification_of_Learning_and_Instruc.html?hl=zh-CN&id=M2Rb9ZtFxccC (accessed October 10, 2021).

Tian, N., Dong, Y., and Wei, Z. (2018). *Impact Analysis of Different Types of Online Games Webcast on the Originality of Game Graphics*. Available Online at: https://webofproceedings.org/proceedings_series/article/artId/3463.html (accessed October 9, 2021).

Tokac, U., Novak, E., and Thompson, C. G. (2019). Effects of game−based learning on students' mathematics achievement: a meta−analysis. *J. Comput. Assist. Learn.* 35, 407–420. doi: 10.1111/jcal.12347

Torrente, J., Moreno-Ger, P., Fernández-Manjón, B., and Sierra, J. L. (2008). "Instructor-oriented authoring tools for educational videogames," in *8th IEEE International Conference on Advanced Learning Technologies (ICALT 2008)*, (Santander: IEEE), 516–518.

Tundjungsari, V. (2020). "Mobile Learning Design Using Gamification for Teaching and Learning in Algorithms and Programming Language," in *The Challenges of the Digital Transformation in Education*, eds M. E. Auer and T. Tsiatsos (Berlin: Springer), 650–661. doi: 10.1007/978-3-030-11932-4_61

Uzelac, A., Gligoric, N., and Krco, S. (2015). A comprehensive study of parameters in physical environment that impact students' focus during lecture using Internet of Things. *Comput. Hum. Behav.* 53, 427–434.

Vate-U-Lan, P. (2015). Transforming Classrooms through Game-Based Learning: a Feasibility Study in a Developing Country. *Int. J. Game Based Learn.* 5, 46–57. doi: 10.4018/ijgbl.2015010104

Villalta, M., Gajardo, I., Nussbaum, M., Andreu, J. J., Echeverría, A., and Plass, J. L. (2011). Design guidelines for classroom multiplayer presential games (CMPG). *Comput. Educ.* 57, 2039–2053.

Vos, N., Van Der Meijden, H., and Denessen, E. (2011). Effects of constructing versus playing an educational game on student motivation and deep learning strategy use. *Comput. Educ.* 56, 127–137.

Wang, M., and Zheng, X. (2020). Using Game-Based Learning to Support Learning Science: a Study with Middle School Students. *Asia Pac. Educ. Res.* 30, 167–176. doi: 10.1007/s40299-020-00523-z

Wang, S., Shi, G., Lu, M., Lin, R., and Yang, J. (2021). Determinants of Active Online Learning in the Smart Learning Environment: an Empirical Study with PLS-SEM. *Sustainability* 13:9923.

Webb, A. W., Bunch, J. C., and Wallace, M. F. G. (2015). Agriscience Teachers' Implementation of Digital Game-based Learning in an Introductory Animal Science Course. *J. Sci. Educ. Technol.* 24, 888–897. doi: 10.1007/s10956-015-9571-7

Wu, W., Hsiao, H., Wu, P., Lin, C., and Huang, S. (2012). Investigating the learning−theory foundations of game−based learning: a meta−analysis. *J. Comput. Assist. Learn.* 28, 265–279.

Xuqing, W. (2007). Second Life and Digital Game-based Learning Environment [J]. *Distance Educ. J.* 4, 76–78.

Yadav, A. K., and Oyelere, S. S. (2020). Contextualized mobile game-based learning application for computing education. *Educ. Inf. Technol.* 26, 2539–2562. doi: 10.1007/s10639-020-10373-3

Yang, J., and Huang, R. (2015). Development and validation of a scale for evaluating technology-rich classroom environment. *J. Comput. Educ.* 2, 145–162.

Yang, J., Pan, H., Zhou, W., and Huang, R. (2018). Evaluation of smart classroom from the perspective of infusing technology into pedagogy. *Smart Learn. Environ.* 5:20. doi: 10.1186/s40561-018-0070-1

Yang, K.-H., and Lu, B.-C. (2021). Towards the successful game-based learning: detection and feedback to misconceptions is the key. *Comput. Educ.* 160:104033. doi: 10.1016/j.compedu.2020.104033

Zhang, W., and Liu, J. (2007). The Examined Thought about Essence and Value of Edugame—Looking at Combination of Education and Game From Game Angle [J]. *Open Educ. Res.* 5, 64–68.

Zimmerling, E., Höllig, C. E., Sandner, P. G., and Welpe, I. M. (2019). Exploring the influence of common game elements on ideation output and motivation. *J. Bus. Res.* 94, 302–312. doi: 10.1016/j.jbusres.2018.02.030

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Psychology

ORIGINAL RESEARCH
published: 11 January 2022
doi: 10.3389/fpsyg.2021.709624

University Credits as a Measure of Teachers' Pre-service and In-Service Training: A Longitudinal Approach Using Swedish Data

Stefan Johansson[1], Åse Hansson[1] and Tarja Alatalo[2]*

[1] Department of Education and Special Education, University of Gothenburg, Gothenburg, Sweden, [2] School of Teacher Education, Dalarna University, Falun, Sweden

OPEN ACCESS

Edited by:
James Ko,
The Education University of Hong Kong, Hong Kong SAR, China

Reviewed by:
Jeffrey M. DeVries, Technical University Dortmund, Germany
Emrah Soykan, Near East University, Cyprus
Balwant Singh, Partap College of Education, India
Seyeoung Chun, Chungnam National University, South Korea

****Correspondence:***
Stefan Johansson
Stefan.johansson@gu.se

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *14 May 2021*
Accepted: *10 December 2021*
Published: *11 January 2022*

Citation:
Johansson S, Hansson Å and Alatalo T (2022) University Credits as a Measure of Teachers' Pre-service and In-Service Training: A Longitudinal Approach Using Swedish Data. Front. Psychol. 12:709624. doi: 10.3389/fpsyg.2021.709624

In this study, we accessed information about the university credits of all teachers born after 1971 in Sweden as a means of ascertaining the development of their subject knowledge. We examined the university credits they earned during pre-service and in-service training. Data comes from registers Gothenburg Educational Longitudinal Database (GOLD) and the teacher register. We linked GOLD to the teacher register in order to describe the knowledge development of teachers in compulsory school 1998–2014. Special focus was on Swedish language and mathematics. Multiple regression and multilevel growth modeling were used as our main methods. Results show an increase in pre-service credits during the time period and more credits in Swedish language than in mathematics. To analyze teachers' in-service training, we followed the development of their university credits over time. Teachers with higher prerequisites in terms of grade point average tended to gain more credits in-service. The study included discussions on ideas and the implications for future research.

Keywords: teacher knowledge, pedagogical content knowledge, university credits, GPA, longitudinal studies, pre-service training, in-service training, professional development

INTRODUCTION

In recent decades, teacher quality has shown itself to be a key issue for schools that significantly affects student achievement (e.g., Nye et al., 2004; Rockoff, 2004; Rivkin et al., 2005; Darling-Hammond, 2016). Despite numerous studies demonstrating the substantial effects of teacher quality, there is little agreement as to how to conceptualize teacher quality or as to which characteristics are effective (e.g., Hanushek and Rivkin, 2010). Darling-Hammond (2021), however, noted some commonalities for successful school-systems. These were characterized by, for example, high-performing student teachers and a master's degree for every teacher, as well as clearly framed standards of what knowledge every teacher should have.

Standards for teacher knowledge largely depart from Shulman's (1986, 1987) work. Shulman developed the well-known theoretical distinction between teachers' content knowledge (CK)

and pedagogical content knowledge (PCK). CK refers to the current knowledge within a subject-domain, while PCK comprises ways of representing and formulating CK that makes it comprehensible to students. PCK represents the blend of content and pedagogy into an understanding of how particular topics and issues are organized, represented, and adapted to the diverse interests and abilities of learners. The necessity of deep subject knowledge has been proposed as foundational to effective teaching. Darling-Hammond (2016), for example, argue that teachers' CK and PCK interact and that they together determine teacher effectiveness. Moreover, teachers need generic competencies concerning, for example, classroom management, lesson planning, and assessment, which is often referred to as general pedagogical knowledge (GPK).

Effective teachers are thus characterized by a complex composite of knowledge and skills, and it is difficult to delineate the components of CK and PCK empirically. Hill et al. (2005), for example, suggested that the CK and PCK in mathematics of primary school teachers could be merged into one unit of knowledge: mathematical knowledge for teaching (MKT). Others have found that CK and PCK are separable and unique dimensions, be they nonetheless correlated (Phelps and Schilling, 2004; Krauss et al., 2008; Baumert et al., 2010). Baumert et al. (2010) showed that despite the high correlation between CK and PCK, CK had lower predictive power for student progress than did PCK. CK was, however, an important prerequisite for PCK. This was confirmed in a similar study by Callingham et al. (2016). Moreover, teachers' CK has been found to relate to the appropriateness in teachers' feedback to students' mathematical ideas and to the accuracy of teachers' mathematical language (e.g., Spillane, 2000; Hill et al., 2008), and also to their implementation of mathematics curriculum materials (e.g., Manouchehri and Goodman, 2000; Sherin, 2002). Mathematics CK is also a significant predictor of pre-school teachers' ability to perceive learning situations and to plan educational actions that foster learning (Dunekacke et al., 2015). Teachers' PCK is shown to relate to their instructional quality (e.g., Kunter et al., 2013), and, for example, to knowledge of students' thinking about mathematics (e.g., Lubinski, 1993).

For the case of reading, primary school students' learning is improving as a result of the training teachers receive pre-service and that prepares them for the profession and the teaching of primary school students (Croninger et al., 2007). Having a master's degree has not shown consistent results for student achievement (Rockoff, 2004; Rivkin et al., 2005; Leigh, 2010). For mathematics in lower-secondary school, however, a master's degree increases the ability of a teacher to boost student achievement (Harris and Sass, 2011), as is also the case for secondary school mathematics teachers (Lachner and Nückles, 2016). Croninger et al. (2007) also show the importance of contextual effects, such as collective expertise at the school level in literacy, which could develop stronger curricular programmes and provide pedagogical support to less qualified colleagues, boosting school-wide disciplinary cognitive gains. As regards mathematics, Seidel and Shavelson (2007) showed in a meta-analysis that the greatest effects of teaching derive from disciplinary components of mathematics teaching obtained during pre-service teacher training, not from the way teaching is organized. Indeed, teachers' course work relating to mathematics is a factor often believed to have positive effects on student achievement (Harris and Sass, 2011). For language teachers, course work in reading (e.g., Clotfelter et al., 2007) and language (e.g., Adams and Lowey, 2007; Hart et al., 2009; Lamb, 2010) is also crucial. However, Monk (1994) showed that the effects of teachers' disciplinary course work in mathematics and science are less significant than additional training in pedagogy.

Besides the pre-service training, teachers' CK and PCK can be developed by in-service training (Kleickmann et al., 2013). Several studies suggest that these aspects of teacher knowledge can improve by way of further training (e.g., Copur-Gencturk, 2015; Kennedy, 2016). To invest in professional development activities may be beneficial for student achievement but also for teachers' self-efficacy and job satisfaction (Boeskens et al., 2020). How to design effective in-service training is, however, not clear. To be effective, research has suggested that the training should focus on CK and be designed as collective and intense participation programmes (Kennedy, 2016). If the programmes focus only teachers' pedagogical knowledge, however, research has documented null effects on student achievement (e.g., Dash et al., 2012). As with mathematics, teachers' knowledge about reading improves through intensive, extended programmes of professional development in literacy (McCutchen et al., 2002; Moats and Foorman, 2003; Carlisle et al., 2011). While several studies show the positive effects of further training, it should be noted that research is inconsistent about how to best develop teachers' knowledge. Some researchers have proposed that the pedagogical part of teacher knowledge is difficult to develop through education and that these skills are better developed by way of work experience (e.g., Hanushek, 2011; Chingos and Peterson, 2013).

Naturally, there are other factors besides professional development programmes that can influence the effectiveness and likelihood to participate in in-service training. Teachers' background, for example, their prior knowledge and teaching experience could influence both the outcome and amount of in-service training teachers enroll in. The extent to which teachers undergo in-service training may also depend on if it is voluntarily or not—some initiatives are school-based and mandatory, whereas others are teachers' own initiatives. The self-determination theory (SDT) may be useful in explaining both why individuals enroll in teacher training as well as why they take in-service training. SDT concerns the motives that steer individual's choices and describes how and why some individuals are more proactive and engaged (Ryan and Deci, 2000). Some of the teachers are likely more self-determined than others; they are driven by intrinsic motivation and strive for competence and autonomy, which indeed are central parts of the SDT. The role of teachers' prior knowledge might influence both the likelihood of enrolling in in-service training and its outcome—especially if training is taken on voluntary basis. In the best of worlds, teachers with less knowledge would enroll in further training to a higher extent than teachers with more knowledge. In that way, successful in-service training could compensate for some teachers' lower prior knowledge and weaker subject knowledge.

This may be a particularly pertinent issue in Sweden where the prior knowledge of new teachers is shown to have decreased in recent years (Bertilsson, 2014; Alatalo et al., 2021). Since the mid-1990's, Sweden experienced extensive school reforms: for example, new curricula, new teacher education, and decentralized governance of schools.

Aim

Teachers develop their knowledge in two major ways: By pre-service training and in-service activities including reflection on teaching experiences as well as professional development which could both be out of school and/or school embedded. Pre-service and in-service activities develops teacher quality and this in turn affects student achievement. However, it has been difficult to establish any significant effects of length of pre-service training (Rivkin et al., 2005) or in-service training (Kennedy, 2016). This is due to a number of reasons, for example lack of precision in how pre-service and in-service activities were operationalized. In this study, we exploit a unique dataset comprising every university credit that teachers earned during both their pre-service and in-service years. We were able to examine every course that teachers took and when they took it, and we were also able to categorize these courses into mathematics and Swedish language domains. However, it was beyond the scope of this study to distinguish between CK and PCK aspects due to the large variation in terminology.

The main aim of the study is rather explorative: We investigate the possibilities to operationalize a more precise measure of teachers' pre-service and in-service training than has been provided in much of the previous research. We will shed light on the development of teachers' pre-service and in-service training in terms of their university credits in Sweden from 1998 to 2014, with special focus on mathematics and literacy. We hypothesize a non-linear increase for in-service training because we anticipate that teachers do not participate in these activities the first years in the profession to high extent. We also hypothesize that teachers with higher own final grades (GPA score) enroll to higher degree in further training and that no compensatory function can be observed from in-service training. Furthermore, we investigate differences relating to the Grade taught: primary (Grades 1–6)/secondary (Grades 7–9). We begin our investigation with determining the number of pre-service credits for teachers during the period 1998–2014.

Specific research questions were:

(1) To what extent did Swedish teachers enroll in university based pre-service and in-service training during 1998–2014?
(2) How did teachers' own GPA influence the likelihood of enrolling in in-service training?
(3) Were there any differences in the amount of in-service training for different Grade levels?

MATERIALS AND METHODS

To answer our research questions, we used data from the Swedish teacher register provided by Statistics Sweden. In this data, the complete population of teachers in Swedish schools is present, including detailed information about their position (e.g., the disciplinary subject they teach), their teacher education, and their certification status. In this study, we linked data from The Gothenburg Educational Longitudinal Database (GOLD), which stores other information about all individuals born after 1971 in the teacher register data. Data from 1998 to 2014 was used in the current study. The credits for the university courses were first registered in the GOLD database in 1993–1994, but relatively few of the GOLD teachers were at that time enrolled in higher education. Since there was also a lag in the registration of credits, we selected 1998 as a starting point for analyses using university credits. The larger research project in which our study was conducted had available data on university credits up to 2014. The registers are presented below in more detail.

The Teacher Register and GOLD

The teacher register forms part of the national follow-up system for the school sector of the Swedish National Agency for Education. Its purpose is to provide a comprehensive picture of school activities as well as support for follow-up and evaluation at the national and regional level. Information is collected annually for all school staff with educational duties (for example, teachers, assistant teachers, pre-school teachers, recreation instructors, school leaders, and study and career counsellors). In the current study we make use of those having a position as teacher. The information about teachers in the register is most often provided by the principal. Data has been collected since the late 1970s, and the structure and variables of the register have changed over the years.

A unique component of the teacher register data is that it is stored by personal identification number, which facilitates a link between the teacher register and the national database GOLD, which also uses the personal identification number system. GOLD contains data for the complete population in Sweden born after 1971. Data on individuals is stored from the year a person is 16 years old and is updated yearly. As such, both the teacher register and GOLD contain longitudinal information at the individual level, where certain characteristics are fixed whereas others vary. For each individual there is comprehensive information about family background, school achievement, employment, and income, among other things. When we merge GOLD with teacher register data, we are able to investigate teachers' pre-service and in-service training in a detailed manner. Since GOLD includes individuals born after 1971, we did not analyze all teachers in compulsory school. The number of GOLD teachers increases over time. For 1998, we mainly looked at young teachers because those born in 1972 are 26 years old in 1998. In 1998, the data on some 5,700 teachers was available in GOLD. In 2014, we analyzed approximately 50% of all teachers in Sweden, at which time the number of GOLD teachers was about 42000.

Variables

Our analysis focuses on disciplinary course work as the variable of interest. This output variable is described in the following. In

addition, we used several covariates to shed light on development over time for different teacher groups. These include: subject (focus on mathematics/Swedish language), Grade [primary (Grades 1–6)/secondary (Grades 7–9)], teachers' own final grades (GPA) from compulsory school (Grade 9). Because grades have been subject to grade inflation, every individual was assigned a percentile transformed grade. Percentiles were computed for each cohort separately, so for each cohort the GPA had a mean of around 50 and a SD of around 28.

We use the total number of teachers' university credits each year as our outcome variable to measure length of pre-service training. A national credit system is used by Swedish universities to show the scope of a course or study program where one week of full-time studies (40 h) corresponds to 1.5 higher education credit. Each semester is 20 weeks long, during which a student is expected to take 30 credits' worth of courses. One Swedish credit is equal to one ECTS credit. The European Community Course Credit Transfer System, was developed by the Commission of the European Communities (ETCS) in order to provide common procedures to guarantee academic recognition of studies abroad. It provides a way of measuring and comparing learning achievements, and transferring them from one institution to another.

Gothenburg Educational Longitudinal Database includes every university course taken by individuals born after 1971. We used the total number of credits held by teachers as well as their credits in mathematics and Swedish language. As we were particularly interested in courses concerning mathematics and Swedish language, we dummy-coded courses that indicated CK or PCK in mathematics and Swedish language and used these as explanatory variables in the analyses.

Analytical Methods

This study relies on several statistical methods. In order to describe the general trends of teachers' pre-service training, we used multiple regression and descriptive analysis methods. We studied the complete GOLD population of teachers as well as the newly recruited teachers each year.

To gain insight into the development of teachers' in-service training, we carried out multilevel growth modeling using the mixed model for repeated measures provided in SPSS 26.0. The design implied that measurements were taken for different individuals at different points in time (teachers signed up for in-service training at different occasions during their career). Such complex designs can be effectively examined using multilevel modeling techniques (Hox, 2002). The basic model for investigating in-service trajectories is presented in Equation 1.The longitudinal analyses allowed us to investigate more deeply how in-service training changed over time as well as if the credit change was the same for teachers with different characteristics. All analyses were carried out using SPSS 26.0.

RESULTS

The research questions concern teachers' university based in-service training and what influences the development of the number of credits. Firstly, we will explore the development of credits taken pre-service in order to provide a comprehensive picture of how the length of teachers' education has varied across the last decades.

Pre-service Course Work

To examine the development of credits in detail, we conducted regression analysis. Because the number of GOLD teachers and the length of teacher education has increased over time, we controlled for time and teachers' age in the regression. We focused on teachers who had an education, either complete or in part. Thereby we selected teachers who had earned at least 60 credits before being employed as teachers. This makes the comparison of total number of credits less sensitive to variation in the proportion of uncertified teachers, which has varied greatly across years. Uncertified teachers that are employed on temporary basis sometimes do not hold any credits from the university.

We began by including only the year in a simple regression in Model 1. The results presented in **Table 1** showed an increase in the number of credits over time, which is not surprising given that teacher education became somewhat longer for teachers who underwent training after 2001. We also included a quadratic time variable in Model 2 to account for any curvilinear trends. Introducing a quadratic year term rendered a model with higher explained variance (R^2), and it was therefore retained in further models. The quadratic term suggests that the increase in credits deaccelerates and flattens out by the end of the time period. We noted that when year was kept constant, results showed older teachers to have somewhat fewer credits. In Model 3, the number of credits to decrease by 0.5 for teachers' age. However, the number of credits increased over time, regardless of teachers' age. In other words, a 26-year-old in 2014 had more credits than a 26-year-old in 1998. Furthermore, we introduced teachers' GPA and the grade they are working in, primary or secondary school, as dummy variables (primary/secondary). Model 4 demonstrates that secondary school teachers have on average 50 credits more than primary school teachers. Secondary school teachers have generally higher GPA than primary school teachers, one reason for this being higher demands for admission to teacher education for secondary school teachers. However, keeping the grades constant, we noted a higher production of credits for those who had a high GPA. In the final step, we introduced three dummies in the model for GPA. We divided the continuous GPA-variable into four categories based on the standard deviation. The reference category was those with a low GPA, i.e., more than 1SD below the GOLD teacher population mean. The results suggest that for those who have more than 1SD above the average GPA (GPA4), the number of credits is nearly 11 points higher, regardless of the grade taught.

Moreover, we noted that Swedish language teachers have more credits than mathematics teachers, which is interesting considering the increased focus on measures to develop mathematics teaching in recent years. We therefore explored the increase in the number of credits in mathematics and Swedish language by way of a detailed categorization of credits into mathematics and Swedish language. **Figure 1** sheds light on

TABLE 1 | Linear regression for the development of teachers' university credits.

Model		Unstandardized coefficients		Standardized coefficients	*t*	*p*
		B	SE	Beta		
1	Intercept	225.452	0.292		770.825	0.000
	Year	2.667	0.028	0.183	96.861	0.000
2	Intercept	208.632	0.472		442.281	0.000
	Year	7.759	0.116	0.533	67.119	0.000
	Year _quad	−0.287	0.006	−0.360	−45.342	0.000
3	Intercept	221.191	1.009		219.243	0.000
	Year	8.087	0.118	0.556	68.602	0.000
	Year _quad	−0.285	0.006	−0.357	−44.983	0.000
	Age	−0.506	0.036	−0.037	−14.081	0.000
4	Intercept	208.204	0.930		223.977	0.000
	Year	7.291	0.108	0.501	67.221	0.000
	Year _quad	−0.251	0.006	−0.316	−43.180	0.000
	Age	−0.581	0.033	−0.042	−17.577	0.000
	Secondary (1)	50.937	0.229	0.385	221.952	0.000
5	Intercept	200.849	0.973		206.503	0.000
	Year	7.566	0.109	0.521	69.538	0.000
	Year _quad	−0.260	0.006	−0.327	−44.615	0.000
	Age	−0.615	0.033	−0.045	−18.621	0.000
	Secondary (1)	50.574	0.230	0.382	220.262	0.000
	GPA4	10.760	0.389	0.065	27.683	0.000
	GPA3	8.010	0.337	0.061	23.800	0.000
	GPA2	6.208	0.363	0.042	17.112	0.000

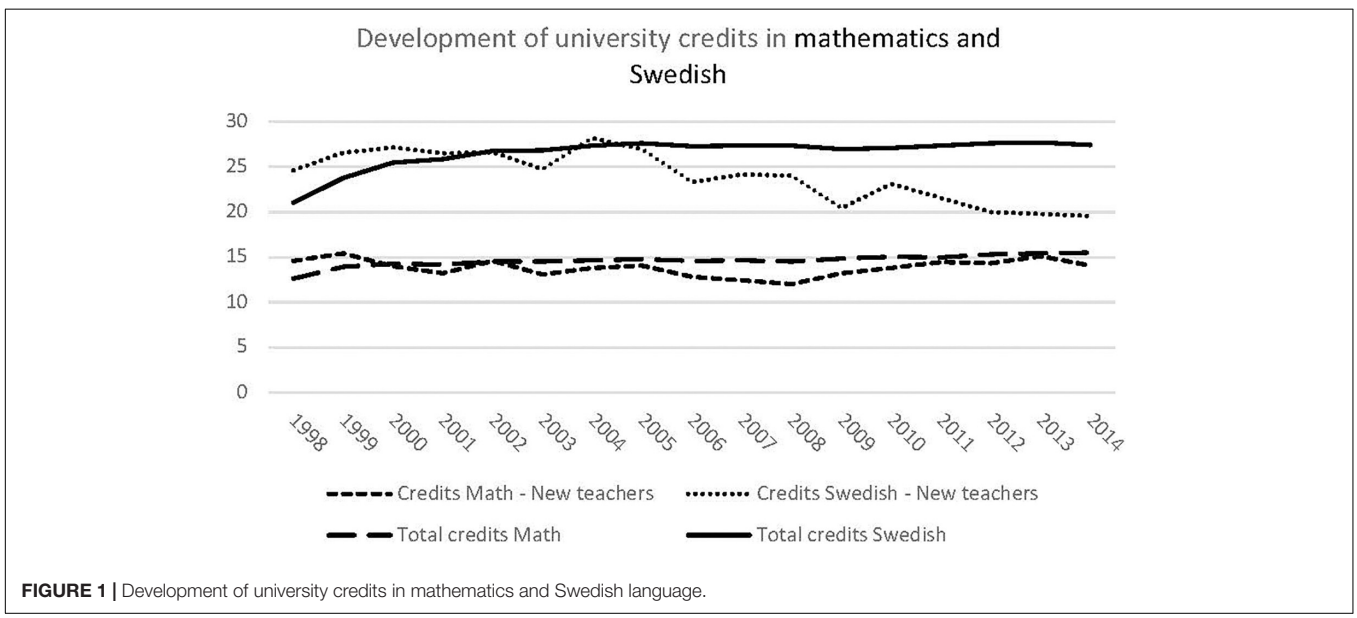


FIGURE 1 | Development of university credits in mathematics and Swedish language.

the number of credits in mathematics and Swedish language, respectively. It may be noted that the proportion of credits is fairly low for both subjects. This is because many teachers in the data set teach subjects other than mathematics and Swedish, and thus naturally do not study these subjects in their pre-service training. More studies are generally devoted to Swedish language than mathematics, possibly because more subjects include training in Swedish than in mathematics. However, for the new teachers, the number of mathematics credits has increased in recent years, while the number of Swedish credits has decreased.

In-Service Course Work

In the following, we will analyze how in-service training developed between 1998 and 2014 by selecting a group of teachers that worked for several years.

We considered university credits that teachers earned after their teacher certification as *in-service training*. In the analyses, we selected teachers who were certified and who received their teaching certification in 1998, meaning they were typically aged between 22 and 26 when they started their teaching career, and who were born between 1972 and 1976. This allows us to study a homogeneous group longitudinally from the years 1998–2014. We followed this graduation cohort over time and studied the development of in-service training with special focus on the cohort's prior knowledge and differences due to their teacher specialization (grades and subjects taught). We investigated in-service training for about 1,500 teachers that at least had 5 years of experience. In total we had about 17,000 observations. The design was complex because not all teachers started to work at the exact same year and they took in-service training at different occasions during their career.

To shed light on teachers' in-service trajectories, we used SPSS Mixed to perform longitudinal analyses by means of a two-level growth model. At level 1, individuals' successive measurements over time were defined by an individual growth trajectory and random error. We used each teacher's number of credits the year after graduation as our dependent variable. Credits accumulated over time for those who enroll in further training. At the second level, differences in trajectories between groups of individuals can be explored. In a first step, we defined the shape of the teachers' growth trajectories by determining whether the initial intercept and random time slope varied across individuals. Next, since the intercept and growth rates varied across individuals, we introduced a set of predictors at level 2 (Grades, GPA, subject taught) to explain the differences in teachers' initial number of credits and their growth trajectories. To investigate this, we constructed cross-level interactions that involved the effects of level 2 variables on level 1 coefficients. Level 2 variables were Grades (primary/secondary), GPA, and subject—variables that do not vary by individual but across individuals. The level 1 slope coefficient was the growth of teachers' yearly earning of credits. Prior to the analysis, we recoded the year variable to match the starting year and the executive years for each teacher (0 for year 1998 if teachers started to work in 1998). We also defined a quadratic time variable (time × time) to capture any changes (acceleration or deceleration) in the rate of change that might occur over the time period. In Equation 1, the basic model without level 2 predictors is presented:

$$Y_{ti} = \pi_{0i} + \pi_{1i} + \pi_{2i} + \varepsilon_{ti}, \quad (1)$$

where,

$$\pi_{0i} = \beta_{00} + r_{0i},$$
$$\pi_{1i} = \beta_{10} + r_{1i},$$
$$\pi_{2i} = \beta_{20} + r_{2i},$$

Where one individuals *i*'s credits Y at time *t* is predicted by an intercept, π_{0i}, and a linear growth slope, π_{1i} as well as a quadratic growth slope π_{2i} at level 1.The subscript *i* indicates that the model estimates a separate intercept and a separate linear growth slope for each person in the sample. Therefore, each teacher in the sample can have a unique linear growth rate and a unique intercept. Between teachers we can investigate three within-individual coefficients as randomly varying where *r* is the residual for each equation. As teachers' in-service develop in different ways we keep all the random terms (*r*) in the equation. Through the substitution of the equations above, we can arrive at the single-equation model without level 2 predictors.

$$Y_{ti} = \beta_{00} + \beta_{10} + (year)_{ti} + \beta_{20}(year\_quad)_{ti} + r_{0i} + r_{1i} + r_{2i} + \varepsilon_{ti} \quad (2)$$

In order to more deeply explore the growth trajectories for different teachers we constructed a series of two-level growth models. We introduced GPA and grade-level (primary/secondary). Disciplinary subject had no relationship to in-service credits and results for this model were therefore not reported. To explore any differences in growth rates, we computed interaction terms. The final model including level 2 predictors is presented in Equation 3.

$$Y_{ti} = \beta_{00} + \beta_{10} + (year)_{ti} + \beta_{20}(year\_quad)_{ti} + \beta_{01}(GPA)_{ti} + \beta_{02}(Secondary)_{ti} + \beta_{11}(year * GPA)_{i} + \beta_{12}(year * secondary)_{i} + \beta_{21}(year\_quad * GPA)_{i} + \beta_{22}(year\_quad * secondary)_{i}\ r_{0i} + r_{1i} + r_{2i} + \varepsilon_{ti} \quad (3)$$

The results are presented in **Table 2**. To assess the fit for the models we ran we compared AIC and BIC values. From Model 1 we note that the average earned credits at their year 0 is 238. The yearly increase β_{10} is on average about 2.2 credits, however, when the quadratic term β_{20} is added in Model 2, model fit improves and the results suggests a significant acceleration over time. This is reasonable as teachers do not enroll in further training just after graduation but rather after some few years. The random part of the model indicates significant variability in both intercept and growth rate. The numbers of the random part of the model are not straightforward to interpret, however, it should be noted that variability is highly significant in all models although it decreases slightly when we add predictors.

Furthermore, in Model 3, we introduce teachers' own GPA to the model. GPA has a positive significant effect on number of credits ($\beta_{01} = 0.34$) suggesting that teachers with higher GPA have more credits. The intercept ($\beta_{00} = 211.50$) decreased substantially as teachers with a GPA of 0 have a lower predicted number of credits. When we keep the parameters in Model 3 constant, the average yearly growth of credits is no longer significant, however, this might be explained by the interaction term between year and GPA ($\beta_{11} = 0.02$). The interaction suggests that yearly growth rate differ between teachers with a different GPA levels. Teachers with higher GPA tend to make slightly more growth per year as compared with their counterparts with lower GPA. When we graphed (see **Figure 2**) the development for four GPA categories, we observed a sharp increase of credits for teachers that were in the highest GPA category.

In Model 4, we introduced grade-level. The significant estimate $\beta_{02} = 53.49$ suggests that teachers at the secondary level have an estimated starting level of their credits that is more than

TABLE 2 | Parameter estimates for the fourth growth models.

		Parameter	Model 1	Model 2	Model 3	Model 4
Fixed effects, initial status π_{0i}	Intercept	β_{00}	237.92** (1.26)	238.50** (1.25)	211.50** (6.55)	212.60** (5.73)
	Slope	β_{01} (GPA)			0.34** (0.08)	0.16* (0.07)
		β_{02} (Secondary)				53.49** (2.54)
Rate of change, π_{1i}(Year)	Intercept	β_{10} (Year)	2.20** (0.11)	1.83** (0.13)	0.41 (0.70)	−0.40 (0.70)
	Slope	β_{11} (GPA × Year)			0.02* (0.01)	0.02* (0.01)
	Slope	β_{12} (Secondary × Year)				−0.21 (0.31)
Rate of change, π_{2i}(Year_quad)	Intercept	β_{20} (Year_quad)		0.05** (0.01)	0.12* (0.06)	0.12* (0.06)
	Slope	β_{21} (GPA × Year_quad)			−0.001 (0.001)	−0.001 (0.001)
	Slope	β_{22} (Secondary × Year_quad)				0.02 (0.02)
Variance		$Var(r_{0i}) = T_{00}$	2,295.82 (85.81)	2,243.97 (84.09)	2,182.39 (82.28)	1,664.64 (62.99)
		$Var(r_{1i}) = T_{11}$	17.15 (0.70)	20.57 (0.96)	20.09 (0.94)	20.15 (0.94)
		$Var(r_{2i}) = T_{22}$		0.11 (0.007)	0.11 (0.007)	0.11 (0.007)
Goodness-of-fit		AIC	129,904.84	128,163.37	126,342.03	125,958.39
		BIC	130,036.41	128,302.67	126,481.09	126,097.45
		Deviance	129,870.84	128,127.37	126,306.03	125,922.39
		Parameters	20	21	23	27

Unstandardized coefficients, standard errors are within parenthesis.
***p < 0.001.*
**p < 0.05.*

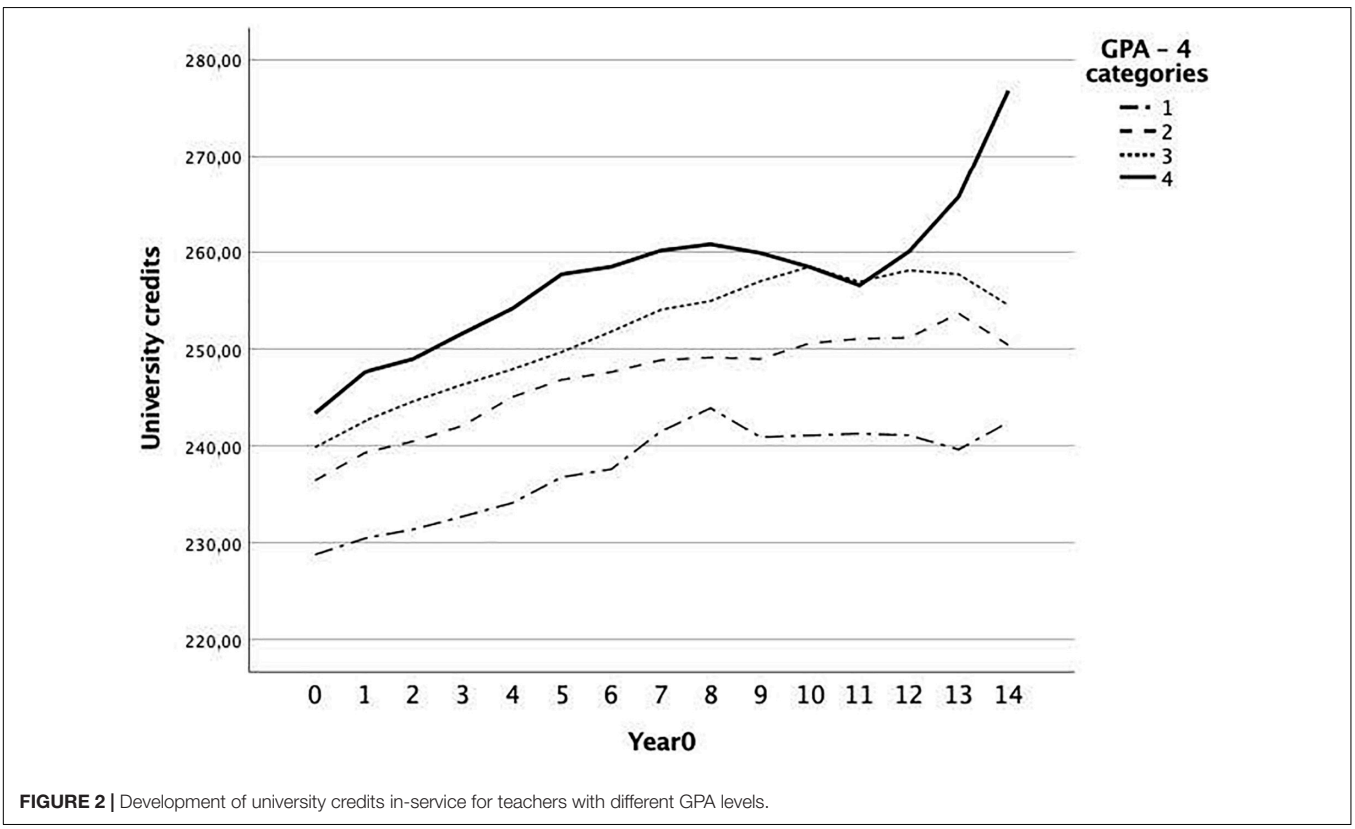


FIGURE 2 | Development of university credits in-service for teachers with different GPA levels.

50 points higher than lower secondary teachers, holding GPA constant. The interaction term is not significant suggesting there were no noteworthy differences in growth per year for teachers at different grades.

It should be noted that teachers enroll in in-service training to a higher extent than what is shown in these findings as we only focused on credits earned at university. Nevertheless, we find no indication that in-service training would work compensatorily in

that teachers with lower prerequisites complete more in-service training than do their counterparts with a higher GPA: in fact, the results show quite the opposite. The growth trajectories seem fairly similar for the different groups of teachers, even though teachers with the higher GPA tended to complete in-service training to a somewhat higher extent.

DISCUSSION

The aim of the study was to investigate teachers' pre-service and in-service training in terms of university credits in Sweden 1998–2014. There was an increase in the number of credits achieved pre-service by teachers in Swedish schools, but we observed a decrease in more recent years. This might have to do with the fact that newly-recruited teachers in recent years begin working before they are fully certified, i.e., they do not have all credits but a fair amount.

Some differences as regards our examination of the number of credits in mathematics and Swedish language we shown. The fact the number of credits pre-service in mathematics has increased in recent years compared with Swedish language may be due to the recent increase in focus on mathematics. Declining results in international mathematics assessments combined with a lack of mathematics teachers may have driven the reforms. While we see less focus on language over time, it is possible that the courses in mathematics include didactical aspects that strongly relate to language. Several studies show that teachers' ability to teach mathematics depends highly on their literacy skills (Adams and Lowey, 2007; Lamb, 2010). In light of the increasing number of immigrant students in Swedish schools, it is crucial that teachers have a large linguistic repertoire so that they can provide effective instruction to those who need it the most.

As expected, we found that teachers at secondary school have more credits than those at primary school. This pattern is constant over time and is a function of the length of education. Regardless, however, of grade, our study suggests that teachers with a higher GPA have more credits. This pattern is constant over time and holds true for both pre-service and in-service training.

International evidence suggests that teachers spend about 11 days per year engaged in professional development activities like courses, workshops and in-service training (Sellen, 2016). However, in the present study, we focused on in-service training in terms of credits achieved at university level. Boeskens et al. (2020) describe an analytical framework for teachers' professional development that might be useful for guiding teachers' in-service training as conceptualized in our study. They describe different dimensions of high-quality in-service training where the *motivation* dimension is particularly relevant for the present study to consider. What shapes teachers' motivation to engage in in-service training at university? To undergo university courses is reasonably an active choice by a motivated teacher rather than an incentive provided by the school principal—who commonly introduce workshop interventions for all teachers at the school. The result that more high-achieving teachers undergo more in-service training might be explained along the lines of the SDT (e.g., Ryan and Deci, 2000). SDT concerns the motives that steer individual's choices and describes how and why some individuals are more proactive and engaged. Some of the teachers are likely more self-determined that others; they are driven by intrinsic motivation and strive for competence and autonomy, which indeed are central parts of the SDT. From this point of view, it is reasonable that the teachers who got higher grades in the end of compulsory school continue to strive to improve and master the challenges they face, for example, undergo in-service training to higher degree than their colleagues who do not have the same prerequisites. We might have expected a compensatory function of in-service training where those with fewer credits and fewer prerequisites (in terms of GPA) would earn more credits over time. Our results can be compared with a study by Smylie (1988) that found that teachers who perceived themselves as being most effective were the same ones most interested in learning new and more effective methods of teaching.

Limitations and Further Research

In this article, we focused on the development of teachers' knowledge in terms of university credits. The number of credits is an excellent quantitative measure that should be further investigated in relation to student achievement. However, it must be noted that this measure neither capture quality of in-service training nor all possible in-service training. Moreover, while we had access to a large sample of teachers, we cannot link teachers to students and cannot investigate the effects of the number of credits on student achievement. An option could be to aggregate teacher and student data to school-level and investigate schools' average level of credits and relate this number to student achievement. We believe that future research should investigate this possibility. It would have been interesting to shed light on the credits in courses with focus on CK and PCK, respectively, but the large number of courses and teachers made this impossible within the scope of this article. For further research, however, we believe that a smaller sample of teachers could be selected for in-depth analysis, where their courses could be investigated in detail to highlight the prevalence of CK and PCK credits. In line with Shulman's (1987) theory, teachers' CK and PCK may improve the reading success of primary school students (Croninger et al., 2007), and to boost students' mathematics in secondary school (Harris and Sass, 2011; Lachner and Nückles, 2016).

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: The data is derived from different registers with census data. Registers are typically accessible for researchers. All data was derived from Statistics Sweden (SCB) and analyzed within MONA (Microdata Online Access) which is Statistics Sweden's platform for access to microdata. In MONA, users process data online without the microdata ever leaving Statistics Sweden. We combined micro level data in MONA from the

teacher register and the Gothenburg Educational Longitudinal Database (GOLD). Requests to access these datasets should be directed to SJ, Stefan.johansson@gu.se.

AUTHOR CONTRIBUTIONS

SJ was the lead author and conducted data analysis and was responsible for writing up the final version of the manuscript. ÅH and TA made substantial contributions in conceptualizing the research and contributed to all parts of the study. All authors contributed to the article and approved the submitted version.

FUNDING

This study was a part of a project which was financially supported by the Swedish Research Council [Grant Number 2016-03636].

REFERENCES

Adams, T. L., and Lowey, R. M. (2007). An analysis of children's strategies for reading mathematics. *Read. Write. Quart.* 23, 161–177. doi: 10.1080/10573560601158479

Alatalo, T., Hansson, Å, and Johansson, S. (2021). Teachers' academic achievement: evidence from Swedish longitudinal register data. *Eur. J. Teacher Educ.* doi: 10.1080/02619768.2021.1962281

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *Am. Educ. Res. J.* 47, 133–180. doi: 10.3102/0002831209345157

Bertilsson, E. (2014). *Skollärare: Rekrytering till Utbildning Och Yrke 1977-2009 [School Teachers: Educational and Professional Recruitment 1977–2009]*. Uppsala: Acta Universitatis Upsaliensis.

Boeskens, L., Nusche, D., and Makito, Y. (2020). *Policies to Support Teachers' Continuing Professional Learning: A Conceptual Framework and Mapping of OECD Data, OECD Education Working Papers 235*. Paris: OECD Publishing.

Callingham, R., Carmichael, C., and Watson, J. M. (2016). Explaining student achievement: the influence of teachers' pedagogical content knowledge in statistics. *Int. J. Sci. Math. Educ.* 14, 1339–1357. doi: 10.1007/s10763-015-9653-2

Carlisle, J. F., Cortina, K. S., and Katz, L. A. (2011). First-grade teachers' response to three models of professional development in reading. *Read. Write. Quart.* 27, 212–238. doi: 10.1080/10573569.2011.560482

Chingos, M. M., and Peterson, P. E. (2013). It's easier to pick a good teacher than to train one: familiar and new results on the correlates of teacher effectiveness. *Econ. Educ. Rev.* 30, 449–465.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2007). Teacher credentials and student achievement: longitudinal analysis with student fixed effects. *Econ. Educ. Rev.* 6, 673–682. doi: 10.1016/j.econedurev.2007.10.002

Copur-Gencturk, Y. (2015). The effects of changes in mathematical knowledge on teaching: a longitudinal study of teachers' knowledge and instruction. *J. Res. Math. Educ.* 46, 280–330. doi: 10.5951/jresematheduc.46.3.0280

Croninger, R. G., Rice, J. K., Rathbun, A., and Nishio, M. (2007). Teacher qualifications and early learning: effects of certification, degree, and experience on first-grade student achievement. *Econ. Educ. Rev.* 26, 312–324. doi: 10.1016/j.econedurev.2005.05.008

Darling-Hammond, L. (2016). Research on teaching and teacher education and its influences on policy and practice. *Educ. Res.* 45, 83–91. doi: 10.3102/0013189x16639597

Darling-Hammond, L. (2021). Defining teaching quality around the world. *Eur. J. Teacher Educ.* 44:3, 295–308. doi: 10.1080/02619768.2021.1919080

Dash, S., de Kramer, R. M., O'Dwyer, L. M., Masters, J., and Russell, M. (2012). Impact of online professional development or teacher quality and student achievement in fifth grade mathematics. *J. Res. Technol. Educ.* 45, 1–26. doi: 10.1080/15391523.2012.10782595

Dunekacke, S., Jenßen, L., and Blömeke, S. (2015). Effects of mathematics content knowledge on pre-school teachers' performance: a video-based assessment of perception and planning abilities in informal learning situations. *Int. J. Sci. Math. Educ.* 13, 267–286. doi: 10.1007/s10763-014-9596-z

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Econ. Educ. Rev.* 30, 466–479.

Hanushek, E. A., and Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *Am. Econ. Rev.* 100, 267–271. doi: 10.1257/aer.100.2.267

Harris, D. N., and Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *J. Public Econ.* 95, 798–812. doi: 10.1016/j.jpubeco.2010.11.009

Hart, S. A., Petrill, S. A., Thompson, L. A., and Plomin, R. (2009). The ABCs of math: a genetic analysis of mathematics and its links with reading ability and general cognitive ability. *J. Educ. Psychol.* 101, 388–402. doi: 10.1037/a0015115

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., et al. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: an exploratory study. *Cogn. Instruct.* 26, 430–511. doi: 10.1080/07370000802177235

Hill, H. C., Schilling, S. G., and Ball, D. L. (2005). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School J.* 105, 11–30. doi: 10.1086/428763

Hox, J. (2002). *Multilevel AnalysisTechniques and Applications*. New Jersey: Lawrence Erlbaum Associates.

Kennedy, M. M. (2016). How does professional development improve teaching? *Rev. Educ. Res.* 86, 945–980. doi: 10.3102/0034654315626800

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., et al. (2013). Teachers' content knowledge and pedagogical content knowledge. *J. Teacher Educ.* 64, 90–106. doi: 10.1177/0022487112460398

Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., et al. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *J. Educ. Psychol.* 100, 716–725. doi: 10.1037/0022-0663.100.3.716

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., and Hachfeld, A. (2013). Professional competence of teachers: effects on instructional quality and student development. *J. Educ. Psychol.* 105, 805–820. doi: 10.1037/a0032583

Lachner, A., and Nückles, M. (2016). Tell me why! Content knowledge predicts process-orientation of math researchers' and math teachers' explanations. *Instruct. Sci.* 44, 221–242. doi: 10.1007/s11251-015-9365-6

Lamb, J. H. (2010). "Reading grade levels and mathematics assessment: an analysis of texas assessment items and their reading difficulty. *Math. Educ.* 20, 22–34.

Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Econ. Educ. Rev.* 29, 480–488. doi: 10.1016/j.econedurev.2009.10.010

Lubinski, C. A. (1993). More effective teaching in mathematics. *School Sci. Math.* 93, 198–202. doi: 10.1111/j.1949-8594.1993.tb12224.x

Manouchehri, A., and Goodman, T. (2000). Implementing mathematics reform: the challenge within. *Educ. Stud. Math.* 42, 1–34. doi: 10.1023/a:1004011522216

McCutchen, D., Harry, D. R., Cox, S., Sidman, S., Covill, A. E., and Cunningham, A. E. (2002). Reading teachers' knowledge of children's literature and english phonology. *Ann. Dyslexia* 52, 205–228. doi: 10.1007/s11881-002-0013-x

Moats, L. C., and Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Ann. Dyslexia* 53, 23–45. doi: 10.1007/s11881-003-0003-7

Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Econ. Educ. Rev.* 13, 125–145. doi: 10.1016/0272-7757(94)90003-5

Nye, B., Konstantopoulos, S., and Hedges, L. V. (2004). How large are teacher effects? *Educ. Evalu. Policy Anal.* 26, 237–257. doi: 10.3102/01623737026003237

Phelps, G., and Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *Elementary School J.* 105, 31–48. doi: 10.1086/428764

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica* 73, 417–458. doi: 10.1111/j.1468-0262.2005.00584.x

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: evidence from panel data. *Am. Econ. Rev.* 94, 247–252. doi: 10.1257/0002828041302244

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intristic motivation, social development, and well-being. *Am. Psychol.* 55:68. doi: 10.1037/0003-066x.55.1.68

Seidel, T., and Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: the role of theory and research design in disentangling meta-analysis results. *Rev. Educ. Res.* 77, 454–499. doi: 10.3102/0034654307310317

Sellen, P. (2016). *Teacher Workload and Professional Development in England's Secondary Schools: Insights From TALIS*. London: Education Policy Institute.

Sherin, M. G. (2002). When teaching becomes learning. *Cogn. Instruct.* 20, 119–150. doi: 10.1207/s1532690xci2002_1

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educ. Res.* 15, 4–14. doi: 10.3102/0013189x015002004

Shulman, L. S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educ. Rev.* 57, 1–23. doi: 10.17763/haer.57.1.j463w79r56455411

Smylie, M. A. (1988). The enhancement function of staff development: organizational and psychological antecedents to individual teacher change. *Am. Educ. Res. J.* 25, 1–30.

Spillane, J. P. (2000). A fifth-grade teacher's reconstruction of mathematics and literacy teaching: exploring interactions among identity, learning, and subject matter. *Elementary School J.* 100, 307–330. doi: 10.1086/499644

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

BRIEF RESEARCH REPORT
published: 09 February 2022
doi: 10.3389/fpsyg.2022.811864

Relationship Between Group Work Competencies and Satisfaction With Project-Based Learning Among University Students

Anabel Melguizo-Garín[1], Iván Ruiz-Rodríguez[1], María Angeles Peláez-Fernández[1]*, Javier Salas-Rodríguez[1] and Elena R. Serrano-Ibáñez[2]*

[1]Department of Social Psychology, Social Work, Social Anthropology, and East Asian Studies, University of Málaga, Málaga, Spain, [2]Department of Personality, Assessment, and Psychological Treatment, University of Málaga, Málaga, Spain

There is a growing interest in improving the teaching–learning process at all levels of education, including higher education. In recent years, university institutions have been taking action to renew and modernize the way in which they teach and learn, making the process more dynamic and closer to the current social reality. Competencies such as the ability to work in a team have become essential for the successful implementation of innovative methodologies in which student participation is particularly relevant. Student acceptance is key to the success of any teaching methodology; however, the influence of group work skills on satisfaction with innovative methodologies such as project-based learning (PBL) has not yet been tested among university students. Thus, the objective of this study is to explore the association between group work competencies and satisfaction with PBL. A total sample of 359 students from two Spanish universities participated in the research. Our results reveal that there is a significant and positive relationship between competencies related to group work and satisfaction with PBL. In addition, a multiple regression analysis shows that the competencies "Conception of group work," "Usefulness of group work," "Planning of group work by teachers," and "Group norms" increase satisfaction with the use of the PBL methodology. This work expands our knowledge about the role in increasing students' satisfaction that is played by the ability of college students to work as a team. These findings could also guide teachers interested in new teaching methodologies.

Keywords: group work, competencies, project-based learning, higher education, satisfaction

OPEN ACCESS

Edited by:
James Ko,
The Education University of Hong Kong, Hong Kong SAR, China

Reviewed by:
José Gutiérrez-Pérez,
University of Granada, Spain
Teresa Pozo-Rico,
University of Alicante, Spain

****Correspondence:***
Anabel Melguizo-Garín
anamel@uma.es
María Angeles Peláez-Fernández
mapelaez@uma.es

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *09 November 2021*
Accepted: *17 January 2022*
Published: *09 February 2022*

Citation:
Melguizo-Garín A, Ruiz-Rodríguez I, Peláez-Fernández MA, Salas-Rodríguez J and Serrano-Ibáñez ER (2022) Relationship Between Group Work Competencies and Satisfaction With Project-Based Learning Among University Students. Front. Psychol. 13:811864. doi: 10.3389/fpsyg.2022.811864

INTRODUCTION

In recent decades, we have been subjected to accelerated change due to new technologies, evolving toward the Information and Knowledge Society. This aspect has also had to be integrated into higher education and its teaching–learning processes. Students start from an abundance of information that is accessible and attractive. However, they lack the skills to manage this information and to know how to use this knowledge to solve day-to-day problems. In this sense, higher education must go beyond the knowledge taught, and teach skills. This aspect

has been highlighted by the United Nations Educational, Scientific and Cultural Organization (UNESCO) (1999), which establishes, as the objectives of education, "to be," "to know," "to do," and "to work together."

To respond to this need, higher education has developed new student-centered teaching–learning methodologies such as project-based learning (PBL; Barrett, 2010). PBL is a type of dynamic methodology that is very suitable for improving the teaching–learning processes of students (Chen et al., 2020). PBL is based on collaborative work between people in a group who must self-direct their work to meet an objective (Clark, 2006). Specifically, PBL presents real problems to students, who must solve them by reflecting and investigating. To do this, they must cooperate, integrate the knowledge of each member of the group, and work both alone and as a team (Wiek et al., 2011).

The implementation of PBL in the higher education system is currently gaining special relevance, as it contributes to the development of work competencies that are very appropriate for the challenges currently faced by society (Kolmos, 1996; Kolmos et al., 2021). There have been studies in higher education that have successfully applied PBL in scientific disciplines of different types, such as chemistry (Morales, 2009), analytical chemistry (Belt et al., 2002), biology (Allen and Tanner, 2003; Pantoja and Covarrubias, 2013), physics (Van Kampen et al., 2004), physiology (Mierson, 1998), and earth sciences (Higgs, 2005). PBL has also been used in other areas of higher education, such as English (Kamiskiené and Janulienė, 2006), education (Iglesias, 2002; Imaz, 2015; Toledo-Morales and Sánchez-García, 2018; Granado-Alcón et al., 2020), history (Galindo, 2008), psychology (López-Zafra et al., 2015; Wiggins et al., 2016), and law (Cubero-Truyo and Díaz-Ravn, 2010), where its use as a didactic strategy has led to significant improvements.

Project-based learning integrates academic knowledge with real-world practices, and undergraduates may not only acquire knowledge and skills and use these in the context of a subject, but may also work on their general competencies (e.g., critical thinking or self-awareness; Brassler and Dettmers, 2017).

In PBL, students have a very high degree of participation and involvement (Kirschner et al., 2006). Given that PBL is a dynamic, interactive, and cooperative methodology, the most relevant competencies involved in PBL are social skills (that is, social, cognitive, and emotional competences that facilitate interaction with others) and group/teamwork skills, that is a set of behaviors, attitudes, and knowledge that contribute toward the struggles of a group to achieve specific common objectives (Stout et al., 1996). There are six skills that are essential for an individual to work efficiently in a group/team: communication, adaptability, coordination, interpersonal skills, decision making, and leadership (O'Neil et al., 1999). In the implementation of participatory and innovative methodologies that require student participation, like PBL, it is necessary to focus on students' teamwork skills and social competencies (El-Adaway et al., 2015). Sometimes PBL may not be beneficial, or its implementation may be difficult, precisely because of the high level of student participation required.

The literature has highlighted the positive aspects of this methodology in contrast to traditional teaching classes and work that is excessively directed by the teacher; these aspects include a greater development of competencies associated with scientific work, a greater knowledge transfer, and a greater satisfaction and affinity with the subject among the students.

Student acceptance is a key to the success of any teaching methodology. Although past research has tested the satisfaction of university students with subjects in which a PBL methodology was used, the influence of group work skills on students' satisfaction with PBL has not yet been tested. In addition, most of the studies mentioned have compared a traditional lecture class to a PBL class in a single degree or at most two degrees, which might limit the generalizability of the research to other degrees.

Therefore, the objective of this study was to explore the association between group work competencies and satisfaction with PBL among a sample of students studying for several different degrees, thus contributing to the body of knowledge about PBL methodology in higher education.

MATERIALS AND METHODS

Participants and Procedure

A total of 359 students from the University of Málaga and the University Isabel I de Castilla took part in the study. The sample consisted of students taking five degree courses: Tourism, Social Work, Psychology, Audiovisual Communication, and Labor Relations and Human Resources. A convenience sampling technique was used, since participants were enrolled in subjects taught by some of the authors of the study. The questionnaires were applied after the students had carried out a piece of group work in their degree subject using the PBL methodology. In order to increase the sincerity of the students, questionnaires were handed by an independent collaborator and students were informed that the answers to the questionnaires were anonymous. The participants completed questionnaires along with a sociodemographic data survey. Although the composition of the sample was biased in favor of females, this imbalance was in line with the gender distribution in the degree courses in which the study was carried out. In fact, female representation in the Tourism and Social Work degree courses is 68 and 85%, respectively, and these are the degree courses from which the majority of the sample came. All participants gave their consent to participate in the study.

The questionnaires were applied after the students had carried out a piece of group work in their degree subject using the PBL methodology. The participants completed questionnaires along with a sociodemographic data survey.

Instruments

The *Questionnaire for the Analysis of Cooperation in Higher Education* (ACOES) was applied in order to measure group work competencies in these higher education students (García et al., 2012). The format response is a five-point Likert scale, ranging from 1 ("Strongly disagree") to 5 ("Strongly agree"). The ACOES comprises 49 items grouped into seven dimensions: Conception of group work, which assesses the student's mental

representations and meaning regarding group work (five items; Cronbach α coefficient = .79); Usefulness of group work, aimed at evaluating student's opinion regarding the usefulness of group work in promoting social interactions, independent learning, and future job performance (six items; α = .80); Planning of group work by teachers, which analyses student's opinion concerning the quantity, complexity, coordination, and mentoring of cooperative work carried out by the teacher (four items; α = .73); Criteria for organizing groups, that assess the criteria that students consider relevant to form group's work: academic o personal reasons, composition (homogeneous or heterogeneous in relation to age, sex, or experience), or temporal stability (eight items; α = .54); Group norms, which evaluate if group norms should be designed by the teacher, the students, or negotiated between both agents (nine items; α = .64); Internal functioning of groups, that assess the sequence of tasks carried out by the students before the final product (seven items; α = .78); and Effectiveness of group work, aimed at evaluating the external and internal conditions of the group that generate better performance and production levels (weighting of group work in the final grade, information on the criteria used, discrimination in the evaluation of the various personal contributions, and inclusion of student self-evaluation and peer evaluation; 10 items; α = .82). The mean scores for each dimension of the ACOES were calculated by adding the scores of the items grouped in the dimension and then dividing by the number of items comprising the dimension.

For the purpose of measuring student satisfaction with PBL, and following the guidelines in the literature on PBL assessment (Dochy et al., 2003; Peterson, 2004; Egido et al., 2007), a *Satisfaction with PBL* (SPBL) self-reporting questionnaire was designed *ad hoc*. The scale comprised 17 items with a five-point Likert scale format response from 1 ("Strongly disagree") to 5 ("Strongly agree"). Cronbach's coefficient was .93. Principal component analysis of the SPBL scale was carried out. Bartlett's test of sphericity was statistically significant [χ^2 (136) = 3307.97, $p < .001$], so the factor analysis was applicable to the selected items. The KMO test shows that the strength of the relationships between the items was high (KMO = .931). Since the scale was designed to assess students' general satisfaction with PBL, the decision was made to extract a single factor. The result was a factor with an eigenvalue of 8.24 that explained 48.46% of the variance. All items had factor loadings higher than .3, which is the minimum limit for considering the factor loading to be significant (Floyd and Widaman, 1995). The mean score for the SPBL was calculated by adding the scores of the items and then dividing by the number of items comprising the questionnaire.

Analytical Strategy

The data analysis was carried out with IBM SPSS version 23. Descriptive statistics were obtained for the socio-demographic variables. Correlational analysis was applied in order to test the association between the SPBL and the ACOES dimensions. Furthermore, a multiple regression analysis with an enter method was carried out with the purpose of checking the predictive value of the dimensions of the ACOES over the SPBL.

RESULTS

Descriptive Analysis and Relationship Between Variables

Descriptive statistics of the participants are shown in **Table 1**. A total of 359 students (74.9% female), with an age range of 18–58, participated in the study. The average age was 22.31 years (*SD* = 4.30).

The correlation coefficients are shown in **Table 2**. These show that satisfaction with PBL was significantly associated with all the ACOES dimensions.

Predictive Model of Satisfaction With PBL

Table 3 shows the results of the multiple regression analysis using the ACOES dimensions as predictors. The regression analysis revealed a significant model for satisfaction with PBL

TABLE 1 | Participant sociodemographic information (*N* = 359).

Age	22.31 (4.30)[a]
Gender	% (*N*)
Male	25.1 (90)
Female	74.9 (269)
Degree	
Labor Relations and Human Resources	19.9 (73)
Tourism	42.3 (155)
Psychology	5.2 (19)
Social Work	28.4 (104)
Audiovisual Communication	4.1 (15)

[a]*Mean (SD).*

TABLE 2 | Correlations of Analysis of Cooperation in Higher Education (ACOES) dimensions with Satisfaction with project-based learning (PBL).

	Satisfaction with PBL
Conception of group work	.495**
Usefulness of group work	.556**
Planning of group work by the teachers	.492**
Criteria for organizing groups	.278**
Group norms	.250**
Internal functioning of groups	.363**
Effectiveness of group work	.345**

***p ≤ .01.*

TABLE 3 | Multiple linear regression model of the dimensions of ACOES for the satisfaction with PBL.

	Unstandardised coefficients		**Standardised coefficients**	***t***
	B	***SE***	***β***	
(model)	.844***	.234		3.615
Conception of group work	.125*	.056	.132	2.224
Usefulness of group work	.325***	.061	.318	5.331
Planning of group work by the teachers	.218***	.041	.257	5.352
Group norms	.132**	.047	.118	2.779

**p < .05; **p < .01; ***p < .001.*
R = .630; R^2 = .396; R^2 adjusted = .390; and F = 58.136.

($F = 58.14$, $p < .001$, $R^2 = .40$). The conception of group work ($\beta = .13$, $p < .05$), the usefulness of group work ($\beta = .32$, $p < .001$), the planning of group work by teachers ($\beta = .26$, $p < .001$), and the group norms ($\beta = .12$, $p < .01$) were significant predictors. The other ACOES dimensions were not significant.

DISCUSSION

Project-based learning is currently gaining special relevance in the higher education system. This methodology requires collaborative work, for which several teamwork skills and social competencies of the students are relevant (El-Adaway et al., 2015). However, it has not been analyzed whether these skills could be related to satisfaction with PBL. Thus, the study aim was to explore the relationship between group skills and students' satisfaction with PBL. The results indicate that satisfaction with the implementation of PBL is related to the teamwork skills of the students. Group competencies, such as "Conception of group work," "Usefulness of group work," "Planning of group work by teachers," and "Group norms" increase satisfaction with the use of the PBL methodology. These skills are closely related to the planning and organization of group work and the role played by the teacher in supporting the planning of the work. These results are in line with previous studies that found association between student satisfaction with PBL and individual and group variables including group harmony (Kilgour et al., 2016), highlighting that the student's own abilities to relate and structure the group can be very important for their satisfaction with the process.

The findings of this study have relevant theorethical implications as they relate group skills and student satisfaction with the PBL methodology. The relationship between both variables is a novel aspect in research of higher education, and underline the need to consider group competencies in future research on higher education employing PBL methodology. On the other hand, these findings have practical implications, as they suggest that the quality of higher education could be improved if teachers played an active role in the implementation of the PBL methodology and designed suitable plans for the group work to be carried out in the implementation of PBL. These plans should address aspects like the design of the group work, the objectives and potential usefulness of the group work, and the rules to be followed by the group, since these issues are closely related to the students' satisfaction with this methodology. In addition, these findings can give guidance, to teachers who want to use new methodologies in their teaching, on the prior skills that are necessary if such an experience is to be satisfactory. Additionally, the inclusion in higher education training of plans aimed at enhancing teamwork skills for students and teachers could be a positive strategy, since such skills could be relevant if the experience of participatory methodologies is to be satisfactory.

A limitation of this research is its cross-sectional character; in future research, a longitudinal study with a control group will be carried out in order to discover whether training in teamwork skills improves satisfaction with PBL. The sample is limited in terms of its origin, since the students belonged only to two universities: University of Málaga and University Isabel I de Castilla, but a positive aspect is that they were studying for different degrees, which contributes to the representativeness of the sample. Another positive aspect of this research lies in its practical implications, since knowing which group competencies can improve satisfaction with PBL can be positive and useful information for improving students' experience with this participatory methodology.

It would be interesting for future research on higher education to explore other individual an contextual factors that may influence satisfaction with PBL methodology, such as the academic performance of students, the number of students in class, or the arrangement of the spaces. In order to generalize the results, it would be appropriate to use samples of students of various university degrees and in different levels and courses. Likewise, the research could be extended to teachers, exploring whether there is a relationship between their abilities and satisfaction with the PBL methodology.

In conclusion, the findings of this research provide a novel aspect in the study of the benefits of using PBL-based methodologies in higher education, as they found that some group competencies, such as group norms, conception of group work, planning of group work by teachers, and perceived usefulness of group work enhance students' satisfaction with PBL. These results suggest that in order to improve the quality of higher education, it would be relevant for teachers to incorporate these group competencies in the study plans and didactics of their subjects. To contribute to the successful implementation of PBL in higher education, it would be useful that both teachers and students were trained with specific group tools such as the ability to plan and organize a teamwork.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Research Ethics Committee of the University of Málaga. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AM-G, IR-R, MP-F, JS-R, and ES-I wrote the first draft. AM-G created and organized the study. AM-G, MP-F, JS-R, and ES-I collected the data. IR-R created the database. IR-R and JS-R analyzed the data. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported and funded in part through research projects from the University of Málaga PIE19-004.

REFERENCES

Allen, D., and Tanner, K. (2003). Approaches to cell biology teaching: learning in context-problem-based learning. *Cell Biol. Educ.* 2, 73–81. doi: 10.1187/cbe.03-04-0019

Barrett, T. (2010). The problem-based learning process as finding and being in flow. *Innov. Educ. Teach. Int.* 47, 165–174. doi: 10.1080/14703291003718901

Belt, S. T., Evans, E. H., McCreedy, T., Overton, T. L., and Summerfield, S. (2002). A problem based learning approach to analytical and applied chemistry. *Univ. Chem. Educ.* 6, 65–72.

Brassler, M., and Dettmers, J. (2017). How to enhance interdisciplinary competence—interdisciplinary problem-based learning versus interdisciplinary project-based learning. *Interdiscip. J. Probl.-based Learn.* 11, 1–12. doi: 10.7771/1541-5015.1686

Chen, J., Kolmos, A., and Du, X. (2020). Forms of implementation and challenges of PBL in engineering education: a review of literature. *Eur. J. Eng. Educ.* 46, 1–26. doi: 10.1080/03043797.2020.1718615

Clark, A. M. (2006). Changing classroom practice to include the project approach. *Early Child. Res. Pract.* 8, 150–172.

Cubero-Truyo, A. M., and Díaz-Ravn, N. (2010). El aprendizaje basado en problemas aplicado al derecho tributario: una propuesta para la asignatura "Fiscalidad Internacional". III Jornadas de Investigación e Innovación Docente, 84–91.

Dochy, F., Segers, M., Van den Bossche, P., and Gijbels, D. (2003). Effects of problem-based learning: a metaanalysis. *Learn. Instr.* 13, 533–568. doi: 10.1016/S0959-4752(02)00025-7

Egido, I., Aranda, R., Cerrillo, R., De la Herrán, A., De Miguel, S., Gómez, M., et al. (2007). El aprendizaje basado en problemas como innovación docente en la universidad: Posibilidades y limitaciones. *Educ. Futuro* 16, 85–100.

El-Adaway, I., Pierrakos, O., and Truax, D. (2015). Sustainable construction education using problem-based learning and service learning pedagogies. *J. Prof. Issue Eng. Educ. Pract.* 141:05014002. doi: 10.1061/(ASCE)EI.1943-5541.0000208

Floyd, F. J., and Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychol. Assess.* 7, 286–299. doi: 10.1037/1040-3590.7.3.286

Galindo, G. (2008). Evaluación de una propuesta didáctica para la enseñanza de la historia. *Revista Vasconcelos de Educación* 4, 33–37.

García, M. M., González, I., and Mérida, R. (2012). Validación del cuestionario ACOES. Análisis del trabajo cooperativo en Educación Superior. *Rev. Invest. Educ.* 30, 87–109. doi: 10.6018/rie.30.1.114091

Granado-Alcón, M. D. C., Gómez-Baya, D., Herrera-Gutiérrez, E., Vélez-Toral, M., Alonso-Martín, P., and Martínez-Frutos, M. T. (2020). Project-based learning and the acquisition of competencies and knowledge transfer in higher education. *Sustain. For.* 12:10062. doi: 10.3390/su122310062

Higgs, B. (2005). "The evolution from problem solving to problem-based learning (PBL): a case study in earth sciences at University College Cork," in *Handbook of Enquiry & Problem-Based Learning*. eds. T. Barrett, I. Mac Labhrainm and H. Fallon (Galway: CELT, NUI Galway), 37–43.

Iglesias, J. (2002). El aprendizaje basado en problemas en la formación inicial de docentes. Perspectivas, XXXII, 1–17.

Imaz, J. I. (2015). Aprendizaje Basado en Proyectos en los grados de Pedagogía y Educación Social:"¿Cómo ha cambiado tu ciudad?"/Project Based Learning in the degrees in Pedagogy and Social Education:" How has your city changed?". *Rev. Complut. Educ.* 26, 679–696. doi: 10.5209/rev_RCED.2015.v26.n3.44665

Kamiskiené, L., and Janulienė, A. (2006). Problem-based learning in the academic setting: language teaching issues. *Santalka. Filologija. Edukologija* 14, 79–87.

Kilgour, J. M., Grundy, L., and Monrouxe, L. V. (2016). A rapid review of the factors affecting healthcare students' satisfaction with small-group, active learning methods. *Teach. Learn. Med.* 28, 15–25. doi: 10.1080/10401334.2015.1107484

Kirschner, P., Sweller, J., and Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educ. Psychol.* 41, 75–86. doi: 10.1207/s15326985ep4102_1

Kolmos, A. (1996). Reflections on project work and problem-based learning. *Eur. J. Eng. Educ.* 21, 141–148. doi: 10.1080/03043799608923397

Kolmos, A., Holgaard, J. E., and Clausen, N. R. (2021). Progression of student self-assessed learning outcomes in systemic PBL. *Eur. J. Eng. Educ.* 46, 67–89. doi: 10.1080/03043797.2020.1789070

López-Zafra, E., Rodríguez-Espartal, N., Contreras, L., and Augusto-Landa, J. M. (2015). Evaluación de una experiencia de Aprendizaje Basado en Problemas (ABP) en estudiantes universitarios. *Revista d'innovació docent universitària* 7, 71–80.

Mierson, S. (1998). A problem-based learning course in physiology for undergraduate and graduate basic science students. *Adv. Physiol. Educ.* 275, S16–S27. doi: 10.1152/advances.1998.275.6.S16

Morales, P. (2009). Uso de la metodología de aprendizaje basado en problemas para el aprendizaje del concepto periodicidad química en un curso de química general. *Revista de la Sociedad Química del Perú* 75, 130–139.

O'Neil, H. F., Lee, C., Wang, S., and Mulkey, J. (1999). Final Report for Analysis of Teamwork Skills Questionnaire. Advanced Design Information, Sherman Oaks, CA.

Pantoja, J. C., and Covarrubias, P. (2013). La enseñanza de la biología en el bachillerato a partir del aprendizaje basado en problemas (ABP). *Perfiles Educ.* 35, 93–109. doi: 10.22201/iisue.24486167e.2013.139.35714

Peterson, T. O. (2004). So you're thinking of trying problem based learning? Three critical success factors for implementation. *J. Manag. Educ.* 28, 630–647. doi: 10.1177/1052562904267543

Stout, R. J., Cannon-Bowers, J. A., and Salas, E. (1996). *The Role of Shared Mental Models in Developing Team Situational Awareness: Implications for Training. In Situational Awareness.* London, England: Routledge, 287–318.

Toledo-Morales, P., and Sánchez-García, J. M. (2018). Project-based learning: A university experience. *Prof. Rev. Curricul. For. Prof.* 22, 429–449.

United Nations Educational, Scientific and Cultural Organization (UNESCO) (1999). *Learning: The Treasure within: Report to UNESCO of the International Commission on Education for the Twenty-First Century*. Paris: Unesco Pub

Van Kampen, P., Banahan, C., Kelly, M., McLoughlin, E., and O'Leary, E. (2004). American journal teaching a single physics module through problem based learning in a lecture-based curriculum. *Am. J. Phys.* 72, 829–834. doi: 10.1119/1.1645280

Wiek, A., Withycombe, L., and Redman, C. L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Sustain. Sci.* 6, 203–218. doi: 10.1007/s11625-011-0132-6

Wiggins, S., Chiriac, E. H., Abbad, G. L., Pauli, R., and Worrell, M. (2016). Ask not only "what can problem-based learning do for psychology?" but "what can psychology do for problem-based learning?" A review of the relevance of problem-based learning for psychology teaching and research. *Psychol. Learn. Teach.* 15, 136–154. doi: 10.1177/1475725716643270

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Psychology

ORIGINAL RESEARCH
published: 18 February 2022
doi: 10.3389/fpsyg.2021.748710

Teaching Facilitation of Family Participation in Educational Institutions

María Ángeles Gomariz[1], Joaquín Parra[1], María Paz García-Sanz[1] and María Ángeles Hernández-Prados[2]*

[1] Department Research Methods and Diagnostic in Education, Faculty of Education, University of Murcia, Murcia, Spain,
[2] Department Theory and History of Education, Faculty of Education, University of Murcia, Murcia, Spain

The participation of families in schools where their children study is a recurring research topic. This field tends to address the perception of parents or teaching staff. This work is novel in that it considers what teachers, and not families, do to facilitate this participation. The purpose of this work has been to contrast a theoretical model with a multidimensional questionnaire designed to obtain information on the assistance provided by teachers to improve parental involvement in schools. It will allow us to lay the foundations for the content necessary for the initial and permanent training of teachers. Then, an initial questionnaire was created and, after being subjected to expert judgment, it was applied to 225 Spanish teachers, using a quantitative and a non-experimental methodology. After calculating the exploratory and confirmatory factor analysis and applying the structural equation model, a questionnaire (QFIS-TP) was obtained that had satisfactory construct validity and reliability.

Keywords: parent participation, teacher, questionnaire, validity, reliability

OPEN ACCESS

Edited by:
James Ko,
The Education University of Hong Kong SAR, China

Reviewed by:
José Gutiérrez-Pérez,
University of Granada, Spain
Kayi Ntinda,
University of Eswatini, Eswatini

****Correspondence:***
María Ángeles Gomariz
magovi@um.es

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *28 July 2021*
Accepted: *10 December 2021*
Published: *18 February 2022*

Citation:
Gomariz MÁ, Parra J, García-Sanz MP and Hernández-Prados MÁ (2022) Teaching Facilitation of Family Participation in Educational Institutions. Front. Psychol. 12:748710. doi: 10.3389/fpsyg.2021.748710

INTRODUCTION

The socioeconomic development of a country is closely tied to its education and ability to generate scientific-technological knowledge and innovations (Spanish Strategy of Science and Technology and of Innovation [EECTI], 2012). Improved educational quality depends not only on economic issues (Organisation for Economic Co-operation and Development [OECD], 2012, 2014), but also on the modernization of systems, through the development of associations and unions for innovation, to capitalize on existing resources (Europe Strategy 2020, 2015). It involves structural and functional changes that affect the school community, requesting the promotion of communication and participatory channels, especially between families and teachers.

Most countries in the European Union have taken on the challenge of promoting school democratization, reassessing the channels of family participation (European Commission, 2005) as a useful measure in response to scholastic failure and student dropout (Organisation for Economic Co-operation and Development [OECD], 2011), and opting for assessment processes that collect student and parent opinions on their educational implication (Maxwell and Staring, 2018). Ultimately, it has been recognized that parents are critical and demanding consumers of

education who can contribute to its improvement through collaborative efforts, or to its decline through their passivity or destructive criticism (European Commission, 2000). In this dichotomy, the prevailing position depends on how family participation is understood, as well as the dynamics promoted by the teaching staff. So, the decision to participate depends on the feeling that one is welcomed by the school, children, and teachers, who utilize various means to express to parents that their participation is necessary and useful (Walker et al., 2005). But, in what content should teachers be trained to promote the participation of families in school?

In addition to the knowledge of degrees of family participation, the complexity of the question requires the creation of a theoretically solid instrument that allows measuring how teachers promote this participation. Although necessary, few studies have analyzed teachers as mediators or promoters of this participation (Erdener, 2016; Gomariz et al., 2017; Yulianti et al., 2019). This is the focal point of this work.

Teachers as Promoters of Family Participation

The main research question guiding the bibliographic review was focused on the role teachers play in family participation in the educational institutions. The lack of scientific research on this topic prevents the identification of the teacher's role in facilitating participation, limiting the search to relevant theoretical models, reviews, and meta-analyses from the past two decades. So, various theoretical models of family participation have been found. The model of Epstein (2011), which distinguishes between six modalities of family participation (parenting, communicating, volunteering, learning at home, decision making and collaborating with community); that of Vogels (2002), which considers four family profiles related to scholastic education (consumers, clients, participants and partners); the renewed model of Hoover-Dempsey and Sandler (Walker et al., 2005), focusing on the aspects that explain why parents get involved and how their participation improves the scholastic performance of their children (involvement decision, forms, influence, mediating variables and student outcomes); Hornby (2011), which includes eight types of family participation (communication, liaison, education, support, information, collaboration, resource, and policy) and that of Garbacz et al. (2019), which considers five participation dimensions (communication, expectations, educational support, community activities, and scholastic assistance). All of these provide a theoretical foundation to form the core of the Integral Model of Family Involvement in School–Teacher Promoter (IMFIS-TP), which we propose and detail below and which includes 7 dimensions. It is a new model that focuses on the teacher figure as the driving force behind specific actions that favor family participation in schools, and not only attributions or good intentions, since expressing the desire and need for increased participation does not contribute to making this participation a reality.

Teacher–Family Communication

The teacher's responsibility in the family communication process has been widely recognized, understanding that parent–teacher tutorials tend to be initiated upon request of the teachers, and less frequently, upon request of the families (Epstein, 2011; Tran, 2014). These communicative meetings offer unique opportunities for the "teacher-promoter" to foster the child's educational support from the home (Dettmers et al., 2019) and the families' sense of belonging to the center (Tran, 2014).

Parents need to understand diverse aspects of their children's' schooling and must feel that they are free to communicate *via* diverse channels and at any time. Therefore, teachers should offer families a wide range of communication channels (Hornby, 2011). The use of agendas, attitude toward participation, the language used, flexibilization of meeting schedules, and the use of the social networks all help to increase family participation, assuming the appropriate use of the same (Erdener, 2016).

School Activities

This dimension focuses on the promotion of collaborative school–family actions in the educational institution, although in combination with activities taking place outside of the school and training activities (community participation and family training, which we will discuss further in this work). Families act as volunteers in a wide variety of activities (school patrolling, student transport, classroom workshops, sporting activities, scholastic camp, fund collection, etc.) which are promoted and coordinated by teachers (Epstein et al., 2019) using distinct participation modalities, such as attendance, collaboration, or involvement in management and decision making (Consejo Escolar del Estado, 2014).

Sense of Belonging

Although research has demonstrated that the sense of belonging to a school is the driving force behind participation (Uslu and Gizir, 2017), few studies have analyzed how teachers contribute to fostering this participation by the families. Reference has been made to the sense of belonging and recognition by the educational community and how this can be promoted through a greater knowledge of the school, trust in the teaching staff, removal of myths and prejudices, meetings with families, satisfaction with the school activities carried out, etc. (Cheung et al., 2017). It is also related to positive communication with teachers (Anderson and Minke, 2007) and involvement in the home (Uslu and Gizir, 2017). Indeed, "Whereas most parents are reluctant to seek the help of professional counselors, they will approach their children's teachers in search of guidance or counseling for the problems that concern them" (Hornby, 2011, p. 37).

Involvement in the Home

The expression "family involvement" refers, on the one hand, to extra-scholastic, cultural, family leisure, and value-transmission activities (Consejo Escolar del Estado, 2014; Caldas and Cornigans, 2015) and, on the other hand, to parent expectations, homework supervision, promoting reading, and reserving the term "participation" to refer to school activities (Castro et al., 2015). Parental support in the home relies on teacher counseling to redirect and encourage the children's studies (Kurtulmus, 2016). It may be offered on an individual or a group manner, and requires the creation of effective, trust-based communicative

processes (Hornby, 2011). All of this contributes to the creation of a sense of belonging to the community.

AMPA (Parent Association) and School Board

The legal model of joint responsibility regulated *via* the parent association (PA) and School Board contributes to improved performance; family participation; educational quality; and knowledge of citizen rights, responsibilities, and duties (Benner et al., 2016). It promotes aspects related to the school and the community, and its success is based in large part on collaboration, recognition, and dissemination of the same by the teacher (García-Sanz et al., 2020). However, although this model is quite extensive and has reached most schools, only a small minority of parents are actively involved (Hornby, 2011), usually as associates and not as managers, be it in the PA management team or as a School Board representative (García-Sanz et al., 2020).

Community Participation

Teachers, in collaboration with the PA and in support of a bidirectional relationship between schools and society, should know how to involve their students' families in cooperation with community/neighborhood organizations. This involves the use of external (libraries, parks, museums, etc.), leisure, volunteer, and solidarity action services, religious events, community associations, or strength–talent training (Epstein et al., 2019; Gahwaji, 2019; Garbacz et al., 2019). According to Severiens et al. (2014), these activities should be included in teacher training curriculums since community projects are more beneficial when they are directed at the entire school population, as opposed to those that are individualized and only from the school.

Family Training

Parental education helps contribute to school improvement. This school-based activity is often managed by PAs and tends to be directed at families, although recently, community training has also been promoted, with parents and teachers learning together (Tran, 2014; Hernández-Prados, 2019). This explains the close association between training, school activities, the PA–School Board, and community participation, since they often rely on municipal training offers or external professionals, especially in the case of public schools (Hernández-Prados et al., 2019).

Using the IMFIS-TP as a foundation, the following research question was proposed: How is it possible to validly and reliably assess the way and the degree to which teachers facilitate family participation in their children's schools? The purpose of the study is to create a multidimensional questionnaire directed at teachers that is based on the IMFIS-TP, which allows us to obtain information regarding the assistance offered by teachers so that parents may participate in the distinct dimensions making up family participation in education institutions. The questionnaire will be very useful to know where to direct the initial and continuous training of teachers.

MATERIALS AND METHODS

Participants

The invited population consisted of 542 teachers from 14 multicultural educational institutes in southeastern Murcia (Spain), where they teach Early Childhood, Primary, and Secondary Education. Of these, 225 teachers agreed to participate, resulting in a 95% confidence level and a 5% sampling error.

Instrument

Based on the IMFIS-TP, an initial questionnaire was created with 11 socio-demographic questions and 91 items on teacher facilitation of family participation, grouped into seven dimensions: "Questionnaire on Family Involvement in School, Teacher-Promoter" (QFIS-TP). After performing an interjudge content validation with 5 university professors (experts in family-school relations and methodology) and the 14 management teams from the participating schools, the instrument kept the 11 socio-demographic questions, but the family participation items were reduced to 74 (**Table 1**).

The scale of items contains 5 degrees, except for item 74, which is dichotomous (Yes/No). In the following dimensions, namely communication with the school, involvement in the home, participation in the PA and school board, training, and all related to school facilitation for. . ., interpretation of the scale is: 1. Never/none; 2. Infrequently/few; 3. Sometimes/somewhat; 4. Often/considerably; 5. A lot/many. In the dimensions of Participation in school activities and Community participation, the scale ranges from: 1. I don't know if they are carried out; 2. I know of them; 3. I attend; 4. I collaborate, participating; 5. I am involved in their organization. Also, those referring to the Sense of belonging dimension are responded to using the following scale: 1. Completely disagree; 2. Disagree; 3. Indifferent; 4. Agree; 5. Completely agree.

Procedure

The validity of the content of the initial QFIS-TP was determined by sending an email to two types of judges (university educators and school management teams), permitting the modification of any items and elimination of others, when considered to be repetitive.

Upon validation, the QFIS-TP was applied to the professors on-line, using the survey platform of the University of Murcia. The research method was quantitative, non-experimental, descriptive and transversal, survey-type. Next, the data were analyzed using the IBM SPSS, v. 24 and IBM SPSS Amos, v.21 programs.

ANALYSIS AND RESULTS

Exploratory Factorial Analysis

To calculate the exploratory factorial analysis (EFA) using the main components extraction and Varimax rotation methods, the SPSS program was used, but first, the Spearman's correlation coefficient was obtained between all items of the instrument to

TABLE 1 | QFIS-TP dimensions and items.

A.Communication with the center

Use an X to mark the frequency or degree to which, in your opinion, the content of each of the questions below arises, based on the indicated scale.

(1) I inform families of the importance of attending teacher-parent conferences when they are requested.

(2) I promote family requests of teacher-parent conferences throughout the school year.

(3) I urge families to attend group meetings of parents with the teacher, which are convened by the teaching staff.

(4) I encourage families, at least, to speak with the teacher in casual meetings at the entrance or exit of the school.

(5) I encourage families to hold meetings with the other teachers in the school, apart from their child's main teacher.

(6) In general terms, the school facilitates communication with the students' parents.

B.Participation in school activities

Indicate how you facilitate family involvement in the following activities, only in the cases in which they are organized by their child's educational institution, based on the indicated scale.

(7) In workshops in the classroom (on reading, crafts, cooking, etc.)

(8) In cultural activities (historic events, musical and ecological topics, traditions, international day of peace, grandparent's day, children's day, women's day, etc.)

(9) In sporting activities (soccer, basketball, judo, karate, etc. camps or exhibitions).

(10 At celebrations (Christmas, Carnival, end of year, etc.).

(11) In outings (to museums, monuments, other institutions, field trips, etc.).

(12) In service activities offered by the school (receiving children in classrooms before the school day begins, library, cafeteria, student transport, etc.)

(13) In work commissions created in the school (co-existence plan, school improvement plan, etc.)

(14) In economic collection commissions for classroom (gifts, costumes, classroom decorations, etc.)

(15) In processes used to assess the school (responding to questionnaires, using the suggestions box, presenting complaints and/or suggestions via PA or individually, etc.)

(16) In general, the school promotes family participation in the activities that it organizes.

C.Sense of belonging

Indicate your degree of agreement or disagreement with the following Statements, based on the indicated scale.

(17) I urge families to identify with the school's values, ideas, attitudes, goals, etc.

(18) I encourage families to feel that they are members of the school, so that they consider it to be their own.

(19) I promote the ideas in families that when a sporting, artistic or cultural team participates in a championship, contest or exhibition, it is their team.

(20) I talk with families to promote their trust in the educational work being carried out by teachers with their children, encouraging support of our decisions.

(21) I encourage families to feel attracted by the collaborative activities or experiences offered by the school.

(22) I help families perceive that their participation in the school makes them a part of the same.

(23) I encourage families to feel welcomed and integrated in the school community as of the onset of their child's schooling.

(24) I do everything possible to ensure that families feel satisfied with the education that the students receive.

(25) I guide families so that they feel free and can express their ideas, concern, suggestions, complaints, etc. in the school.

(26) I encourage families to recommend this school to others with school-aged children.

(27) Generally speaking, the school assists in creating a bond between it and families.

D.Involvement in the home

Mark the frequency or degree to which, in your opinion, the content of the statements presented below takes place, based on the indicated scale.

(28) I guide parents in speaking with their children about what they do in class.

(29) I encourage families to express their trust in their children.

(30) I inform families of their child's class attendance.

(31) I encourage families to take an interest in the educational tasks carried out by their children at home.

(32) I inform families of the importance of their children's study time organization.

(33) I encourage families to offer a good at-home study climate (to encourage studying, offer an appropriate study site without distractions, provide resources for learning, etc.)

(34) I inform families of the need to demonstrate their availability to their child's needs regarding their school work.

(35) I invite parents to congratulate their children when they complete their school work.

(36) I notify families of the importance that their children complete extra-curricular or complementary activities (languages, computer-based, music, dance, sports, academic activities, etc.)

(37) I communicate with families to promote the autonomy and responsibility of their children in their studies (encouraging them to be alert, but not to do their child's work or always be next to them when doing it).

(38) I speak to families so that they oversee the responsible use of computers, mobile phones, etc.

(39) I encourage families to engage in cultural activities (read, go to the cinema, theater, museums, trips, concerts, exhibitions, etc.)

(40) I encourage families to ensure that their children apply what they learn in school to their everyday lives.

(41) In general, the school facilitates parent involvement at home in the educational process of their children.

E.Involvement in the PA and the School Board

Mark the frequency or degree to which, in your opinion, the content of the following statements takes place, based on the indicated scale.

(42) I speak with families so that they are informed with regard to the organization and functioning of the PA.

(43) I encourage families to know the members of the PA Board of Directors.

(44) I help families gain knowledge regarding the activities organized by the PA.

(Continued)

TABLE 1 | (Continued)

(45) I speak with parents so that they are informed and use the municipal bank of books in which the PA participates.
(46) I encourage families to consult information on the PA via websites, the social networks, etc.
(47) I invite families to participate in activities organized by the PA.
(48) I encourage families to form a part of the school's PA Board of Directors.
(49) I inform the PA of the importance that the association represents the interests of all of the school's families.
(50) Generally speaking, the school promotes parent participation in the PA.
(51) I help families to get to know the organization and functioning of the School Board.
(52) I encourage families to get to know their School Board representative.
(53) I invite families to be informed of the decisions made in School Board meetings.
(54) In encourage families to be informed as to the elections process of the School Board (calendar, candidates, voting procedure, etc.).
(55) I encourage families to participate in School Board elections.
(56) I encourage parents to apply for family representative positions in School Board elections.
(57) In general, the school promotes family participation in the School Board.
F.Community participation
Indicate how you facilitate family involvement in the following activities related to the community, based on the indicated scale.
(58) In collection activities (collection of food, clothes, caps, solidarity markets, etc.)
(59) In ecological activities (cleaning of waterways, march in support of the environment, environmental awareness programs, tree planting, etc.)
(60) In activities organized by neighborhood associations (block parties, neighborhood meetings, neighborhood or city needs, presence in the neighborhood councils, etc.)
(61) In solidarity and volunteering activities (assistance to the elderly, the sick, those with limited resources, those who are alone, soup kitchens, etc.).
(62) In activities of the distinct religious communities.
(63) In activities directed at diversity awareness and integration (gender, abilities, cultural, ethnic background, etc.).
(64) In collaboration activities with youth associations promoting healthy leisure and free time.
(65) In general, the school promotes family participation in community activities.
G.Training
Mark the frequency or degree to which, in your opinion, the content of each of the statements below arises, based on the indicated scale.
(66) I guarantee that the families are informed of the training activities directed at the same that are organized by the education institution.
(67) I encourage families to attend training activities organized by the school.
(68) I speak with families to encourage their participation in training activities that are intended for them.
(69) I encourage parents to get involved in the management of training activities for families.
(70) I notify families of the importance of receiving appropriate training with regard to their children's educational process.
(71) I relate the training that is being offered by the school to an improvement in family-school relations.
(72) Generally speaking, the school facilitates participation in training activities intended for families.
(73) I need training in order to better facilitate family involvement in the education of their children.
(74) I am interested in attending training activities to improve family participation in the education of their children.

avoid problems of multicollinearity. No bivariate correlations exceeding 0.85 were found; therefore, it was not necessary to eliminate any item (Kline, 2005). In addition, the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy of 0.928 and the statistical significance of the Bartlett sphericity test (0.000) permit the EFA.

Coinciding with the dimensions of the IMFIS-TP, seven components were determined, having an explained variance of 64.70%. This initial EFA consisted of all of the ordinal items of the questionnaire, except for those related to the teacher's perception of what the school does to facilitate family participation. With these items, a second EFA was created, since it was considered that its content did not specifically address teacher's attitudes and behaviors to collaborate in the improvement of family participation. **Table 2** shows the rotated components matrix, ordered by size, referring to the first EFA.

As seen in **Table 2**, the first factor includes 14 items of the QFIS-TP dimension called involvement in the PA and School Board; the second 13 items are from the involvement in the home dimension; the third includes the 10 items corresponding to the sense of belonging dimension; the fourth consists of the 7 items from the community participation dimension; the fifth includes 9 items regarding the participation in school activities dimension; the sixth is made up of the 7 items from the training dimension; and the seventh and final factor includes the 5 items of the questionnaire regarding the communication with the school dimension.

Based on the results obtained in this EFA, it can be concluded that the QFIS-TP has not experienced any variations with regard to the assignment of items to each dimension in the content validation carried out by the experts and based on the premises of the IMFIS-TP. Therefore, the naming of the instrument's dimensions, as well as that of the items making up each of these, has been maintained with no modifications.

The second EFA, based on items related to the school's facilitation of families to improve the dimensions considered in the QFIS-TP, obtained a mean Kaiser–Meyer–Olkin sampling adequacy of 0.875 and a statistical significance according to the Bartlett's Sphericity test of 0.000. The variance explained with Eigen values >1 was 55.95% (**Table 3**).

As we can observe, the eight items converge in one unique factor called: Facilitation of the school for family participation.

Confirmatory Factorial Analysis

Using the AMOS program, the IMFIS-TP was corroborated. To do so, the construct validity of the QFIS-TP was

TABLE 2 | First EFA: rotated components matrix.

Items	Component						
	1	2	3	4	5	6	7
Q51	**0.862**	0.076	0.089	0.169	0.103	0.167	0.094
Q55	**0.861**	0.032	0.156	0.142	0.144	0.135	0.068
Q54	**0.856**	0.065	0.173	0.151	0.170	0.170	0.063
Q56	**0.843**	0.068	0.132	0.170	0.200	0.174	0.023
Q52	**0.832**	0.063	0.088	0.193	0.207	0.196	0.006
Q48	**0.829**	0.132	0.093	0.165	0.215	0.178	0.019
Q53	**0.798**	0.133	0.142	0.203	0.174	0.214	0.026
Q42	**0.772**	0.241	0.083	0.105	0.195	0.080	0.179
Q43	**0.771**	0.219	0.082	0.163	0.219	0.195	0.004
Q49	**0.768**	0.083	0.133	0.233	0.266	0.075	−0.010
Q47	**0.768**	0.217	0.125	0.170	0.114	0.094	0.143
Q44	**0.754**	0.235	0.100	0.072	0.262	0.134	0.174
Q46	**0.709**	0.270	0.078	0.265	0.223	0.091	0.064
Q45	**0.634**	0.267	0.105	0.187	0.037	0.134	0.236
Q34	0.115	**0.773**	0.146	0.126	0.070	0.148	0.065
Q33	0.169	**0.797**	0.137	0.065	−0.041	0.077	0.127
Q32	0.229	**0.759**	0.166	0.036	0.035	0.102	0.220
Q40	0.086	**0.757**	0.189	0.124	0.196	0.216	−0.044
Q37	0.102	**0.718**	0.177	0.053	0.057	0.090	0.302
Q31	0.153	**0.696**	0.151	−0.075	0.136	0.110	0.071
Q38	0.241	**0.668**	0.164	0.099	0.088	0.025	0.171
Q28	0.054	**0.658**	0.266	0.080	0.315	0.141	0.124
Q29	0.013	**0.608**	0.347	0.107	0.246	0.283	0.072
Q35	0.036	**0.572**	0.151	0.151	0.074	0.382	0.207
Q39	0.264	**0.534**	0.147	0.106	0.299	0.291	−0.065
Q30	0.176	**0.451**	0.157	0.085	0.002	−0.013	0.130
Q36	0.165	**0.354**	0.062	0.219	0.268	0.314	−0.043
Q24	−0.045	0.191	**0.798**	0.010	0.006	0.020	0.010
Q22	0.191	0.134	**0.767**	0.089	0.264	0.129	0.037
Q21	0.131	0.172	**0.764**	0.099	0.281	0.169	0.061
Q23	0.143	0.149	**0.762**	0.164	0.119	0.151	0.109
Q18	0.144	0.213	**0.761**	0.099	0.218	0.103	0.095
Q20	0.055	0.321	**0.707**	−0.008	0.078	0.013	0.092
Q17	0.170	0.332	**0.671**	0.071	0.115	0.137	0.057
Q25	0.212	0.289	**0.661**	0.062	0.113	0.076	0.149
Q26	0.290	0.058	**0.497**	0.216	0.065	0.225	0.049
Q19	0.204	0.026	**0.421**	0.216	0.269	0.322	0.074
Q61	0.242	0.093	0.085	**0.820**	0.136	0.083	0.017
Q59	0.170	0.080	0.041	**0.787**	0.185	0.151	0.096
Q63	0.211	0.106	0.162	**0.778**	0.092	0.196	0.061
Q60	0.328	0.072	0.069	**0.771**	0.232	0.087	0.035
Q62	0.296	0.020	0.079	**0.734**	0.153	0.006	−0.077
Q64	0.338	0.191	0.161	**0.678**	0.207	0.185	0.016
Q58	0.133	0.135	0.140	**0.531**	0.178	0.248	0.268
Q10	0.131	0.110	0.144	0.068	**0.720**	0.115	0.109
Q8	0.310	0.042	0.146	0.099	**0.683**	0.131	0.119
Q9	0.296	−0.003	0.078	0.198	**0.658**	0.108	0.165
Q11	0.218	0.215	0.032	0.262	**0.636**	0.137	−0.168
Q7	0.164	0.059	0.134	0.053	**0.603**	−0.005	0.162
Q12	0.270	0.188	0.104	0.235	**0.589**	0.074	0.034
Q14	0.140	0.089	0.241	0.107	**0.532**	0.028	0.080

(Continued)

TABLE 2 | (Continued)

Items	Component						
	1	2	3	4	5	6	7
Q15	0.214	0.128	0.185	0.169	**0.523**	0.169	0.074
Q13	0.258	0.188	0.165	0.317	**0.507**	0.162	0.027
Q68	0.393	0.262	0.206	0.147	0.141	**0.728**	0.094
Q67	0.336	0.259	0.267	0.170	0.182	**0.697**	0.179
Q69	0.402	0.218	0.104	0.218	0.285	**0.681**	0.075
Q70	0.292	0.281	0.170	0.220	0.183	**0.680**	0.114
Q71	0.324	0.218	0.215	0.207	0.179	**0.646**	0.137
Q66	0.375	0.269	0.197	0.166	0.051	**0.626**	0.227
Q73	0.243	0.152	0.097	0.044	0.044	**0.247**	−0.099
Q3	0.144	0.335	0.188	−0.013	0.158	−0.003	**0.703**
Q1	0.069	0.395	0.042	0.037	0.168	0.131	**0.673**
Q2	0.138	0.283	0.155	0.090	0.161	0.174	**0.617**
Q5	0.330	0.120	0.178	0.322	0.092	0.101	**0.525**
Q4	0.035	0.174	0.073	−0.030	0.275	0.158	**0.284**

Bolded values are factor loadings of the items that make up each factor or component.

ratified through the calculation of a confirmatory factorial analysis (CFA) using the structural equations model. To offer more sense to the theoretical model, the involvement in the PA and the School Board was separated into two dimensions. Similarly, missing values and outliers were eliminated, despite the fact that there was no agreement regarding the removal of the latter (Aguinis et al., 2013). Therefore, cases in which the standardized observable variables exceeded a score of | 3| were excluded (Verdugo et al., 2008).

In accordance with the IMFIS-TP, the correlation between latent and observable variables, the measurement error of the same, and the covariance between the latent variables and measurement errors are represented graphically in **Figure 1**.

The maximum likelihood method was used for model calculation, complying with the univariate normality criteria, having obtained values up to | 2| for asymmetry and up to | 7| for kurtosis (Curran et al., 1996).

Table 4 presents the regression coefficients (factorial loads) between the observable and latent variables, standard error (SE), critical ratio (CR), and the corresponding statistical significance (p). Likewise, it indicates the standardized regression coefficients between the observable and latent variables. It may be observed that all of the pairs are significant, with $\alpha = 0.01$. Likewise, in all cases, the standardized regression coefficients exceed the typical value of the effect size ($r \geq 0.3$), as determined by Cohen (1988).

TABLE 3 | Second EFA: components matrix.

Items	Component
	1
Q27	0.828
Q41	0.804
Q57	0.784
Q50	0.737
Q65	0.729
Q16	0.715
Q72	0.707
Q6	0.667

Table 5 reveals the covariance coefficients between the latent variables and between the measurement errors of the observable variables, the SE, CR, statistical significance (p), and the respective standardized regression coefficients. As seen, the relationship between 84.62% of the pairs is significant at $\alpha = 0.01$, whereas the relationship between the remaining 15.38% is statistically significant at $\alpha = 0.05$. Likewise, the correlation coefficients reach (approximately) or exceed, in 79.49% of the cases, the typical value of the effect size ($r \geq 0.3$) as established by Cohen (1988).

In accordance with Hu and Bentler (1998), various index types have been used to assess the fit of the model: standardized Chi square or the relative Chi square over degrees of freedom (CMIN/DF), included in the measures of goodness of fit; comparative fit index (CFI), framed within the incremental fit measures; and the root mean square error of approximation (RMSEA), included in the absolute goodness of fit measures. **Table 6** indicates the values obtained and those desired, according to the classification established by distinct authors (Hu and Bentler, 1998; Lévy and Varela, 2003; Hair et al., 2008; Cupani, 2012).

With these referents (**Table 6**), it may be affirmed that the IMFIS-TP has acceptable fit indices between the theoretical structure of the model and the empirical results obtained.

Reliability of the Questionnaire

To obtain the reliability of the QFIS-TP, once again, the SPSS program was used. This psychometric property, both globally and by dimensions, was calculated using the

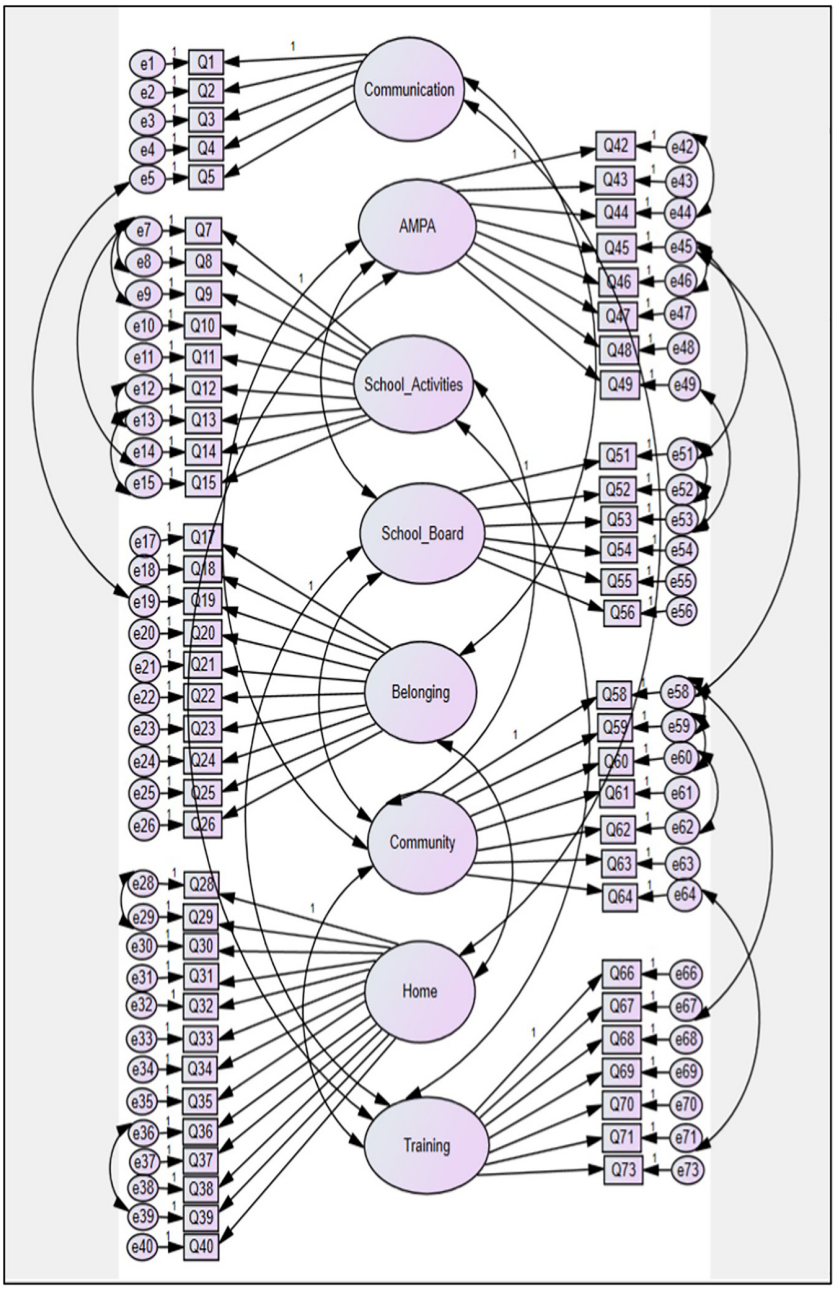


FIGURE 1 | Structural equations model of the QFIS-TP.

Cronbach's alpha dimension (α) and McDonald's omega (Ω). **Table 7** shows the satisfactory indices of internal consistency (DeVellis, 2003), since all of the dimensions exceeded the value of 0.7.

All of the items complied with the corrected item-total correlation; thus it was not necessary to eliminate any of them, either in the QFIS-TP or in the dimensions making up the same. The dimension having the lowest reliability was that of family communication with the school, and the most consistent one was family involvement in the PA and School Board.

DISCUSSION

This study confirms that the psychometric quality of the QFIS-TP, adjusted to the theoretical approach of the IMFIS-TP, is satisfactory. So, the seven dimensions making it up explain a variance of 64.70%, having a high reliability, without modifications in the location of the items with respect to the model. However, the results of the CFA reveal that the octodimensional structure is best, dividing PA and the School Board into two dimensions, given that they

TABLE 4 | Regression coefficients and standardized regression coefficients between observable and latent variables.

Relationship between and latent variables		Regression weights				Standardized regression weights
		Estimate	SE	C.R.	P	Estimate
Q5	Communication	1.312	0.197	6.674	***	0.535
Q4	Communication	1.253	0.253	4.959	***	0.396
Q3	Communication	1.340	0.149	9.000	***	0.771
Q2	Communication	1.410	0.163	8.633	***	0.727
Q1	Communication	1.000				0.697
Q14	School_Activities	1.225	0.238	5.149	***	0.482
Q13	School_Activities	1.528	0.294	5.200	***	0.635
Q12	School_Activities	1.509	0.287	5.255	***	0.650
Q11	School_Activities	1.718	0.320	5.377	***	0.689
Q10	School_Activities	1.218	0.226	5.393	***	0.695
Q9	School_Activities	1.694	0.280	6.045	***	0.732
Q8	School_Activities	1.491	0.232	6.440	***	0.723
Q7	School_Activities	1.000				0.434
Q15	School_Activities	1.225	0.251	4.884	***	0.548
Q66	Training	1.000				0.775
Q67	Training	1.104	0.081	13.548	***	0.857
Q68	Training	1.290	0.092	13.961	***	0.893
Q69	Training	1.321	0.099	13.342	***	0.862
Q70	Training	1.099	0.090	12.248	***	0.806
Q71	Training	1.046	0.089	11.752	***	0.772
Q22	Belonging	1.340	0.126	10.598	***	0.835
Q21	Belonging	1.321	0.122	10.827	***	0.856
Q20	Belonging	0.994	0.115	8.633	***	0.666
Q19	Belonging	0.897	0.152	5.897	***	0.439
Q18	Belonging	1.076	0.113	9.510	***	0.740
Q17	Belonging	1.000				0.683
Q23	Belonging	1.339	0.128	10.439	***	0.821
Q24	Belonging	0.757	0.093	8.121	***	0.624
Q25	Belonging	1.024	0.120	8.535	***	0.658
Q26	Belonging	0.866	0.149	5.812	***	0.439
Q34	Home	1.020	0.108	9.423	***	0.759
Q33	Home	0.894	0.095	9.423	***	0.759
Q32	Home	0.894	0.092	9.700	***	0.786
Q31	Home	0.769	0.093	8.270	***	0.653
Q30	Home	0.733	0.103	7.128	***	0.553
Q29	Home	1.148	0.083	13.907	***	0.760
Q28	Home	1.000				0.663
Q35	Home	0.924	0.105	8.836	***	0.704
Q36	Home	0.941	0.151	6.244	***	0.480
Q37	Home	0.960	0.098	9.757	***	0.791
Q38	Home	1.061	0.122	8.687	***	0.690
Q39	Home	1.106	0.148	7.491	***	0.585
Q40	Home	1.235	0.134	9.226	***	0.740
Q45	AMPA	0.882	0.085	10.341	***	0.651
Q46	AMPA	1.068	0.080	13.393	***	0.786
Q44	AMPA	1.010	0.054	18.805	***	0.834
Q43	AMPA	1.133	0.072	15.718	***	0.866
Q42	AMPA	1.000				0.840
Q47	AMPA	1.188	0.070	16.868	***	0.901
Q48	AMPA	1.329	0.078	17.121	***	0.908

(Continued)

TABLE 4 | (Continued)

Relationship between and latent variables		Regression weights				Standardized regression weights
		Estimate	SE	C.R.	P	Estimate
Q49	AMPA	1.262	0.085	14.842	***	0.837
Q53	School_Board	1.000	0.060	16.611	***	0.855
Q54	School_Board	1.152	0.051	22.418	***	0.974
Q52	School_Board	1.059	0.047	22.763	***	0.887
Q51	School_Board	1.000				0.869
Q55	School_Board	1.177	0.056	20.966	***	0.949
Q56	School_Board	1.191	0.057	20.904	***	0.948
Q61	Community	1.751	0.207	8.447	***	0.840
Q60	Community	1.530	0.188	8.125	***	0.786
Q59	Community	1.554	0.168	9.238	***	0.739
Q62	Community	1.378	0.182	7.563	***	0.701
Q63	Community	1.771	0.209	8.471	***	0.845
Q64	Community	1.611	0.195	8.274	***	0.795
Q58	Community	1.000				0.555
Q73	Training	0.466	0.105	4.445	***	0.323

*** *Statistical significance of the regression between each item and the assigned construct.*

refer to distinct representation organisms (García-Sanz et al., 2020).

The results of the structural equations model have offered covariance between measurement errors of some of the observable variables. However, while this covariance is inevitable, it may be considered appropriate (Landis et al., 2009). In this case, although the reduction of the affected items is similar, they are necessary to confirm the specified theoretical model.

So, based on the results obtained in each of the QFIS-TP subscales, it is evident that the "promoter of communication with families" and "promoter of sense of belonging" roles are theoretical and statistically related (Tran, 2014), and demonstrate a high internal consistency, although covariance has been registered between the measurement error of the variables, encourage meetings with other teachers (Q5), and defense of the school in competitions (Q19). On the other hand, the teacher's role as "promoter of school activities" has produced diverse covariance between measurement errors, suggesting that for teachers, the actions are generalized, categorically hindering their definition. So, we believe that it is necessary to offer more specialized training with this respect, since studies clearly differentiate between classroom (Q7), cultural (Q8), sporting (Q9), and fund collection support (Q14) activities (Hornby, 2011; Epstein et al., 2019). Likewise, although work by commissions (Q13) as a resource may be useful in service (Q12) and school evaluation (Q15) activities, they have specifications of differentiated contents and they operate on different planes. In the first case, in the design of plans and in the following, in implementation and participation. However, Lingard et al. (2014) positively assessed family inclusion in work commissions for educational assessment and restructuring, since they promote active listening and multidirectional dialogue, despite teachers' difficulties in taking advantage of the parents' opinions.

As for "promoting involvement in the home," covariance has been identified in the following measurement errors: Q28–Q29 and Q36–Q39. While positive parent-child relationships are sustained by communication and trust (Q28), this latter should not remain implicit, but rather, must be manifested (Q29) (Ebbert et al., 2019). On the other hand, a confusion between extracurricular (Q36) and cultural (Q39) activities has been confirmed, treating them as synonymous (Ladky and Peterson, 2008), when in fact, the former are academic and individualistic activities, while the latter are collective and are linked to family leisure (Hernández-Prados and Álvarez-Muñoz, 2019).

The subscale to promote the PA is not a generalized action of the teacher consisting of attending meetings (Garbacz et al., 2019), but, on the contrary, it consists of diverse aspects, since helping families getting familiarized with the PA as an organization (Q42) is different from informing them of activities offered to the families (Q44), just like encouraging the use of a banks of books (Q45) and checking the PA website (Q46). On the other hand, according to García-Sanz et al. (2020), the School Board is not as well-known as the PA and requires that teachers promote the knowledge of their representatives (Q52), but also, their organization and functioning (Q51) and the decisions that they make (Q53). Similar items have been proposed by Yulianti et al. (2019). Finally, the relationship between both dimensions reveals itself theoretically (Consejo Escolar del Estado, 2014; Epstein et al., 2019) in the covariance between measurement errors (Q45–Q51 and Q49–Q53), even though each was saturated in different factors.

Seventh, the teacher as "promoter of community participation" is recognized in educational policy, but has

TABLE 5 | Covariance and correlation between latent and observable variables.

Relationship between latent variables and between measurement errors		Covariance				Correlation
		Estimate	SE	C.R.	*P*	Estimate
AMPA	Community	0.264	0.056	4.735	***	0.449
Training	Community	0.233	0.049	4.756	***	0.484
School_Board	Community	0.275	0.062	4.450	***	0.397
School_Activities	Community	0.122	0.035	3.464	***	0.367
Training	AMPA	0.399	0.064	6.209	***	0.587
AMPA	School_Board	0.899	0.111	8.098	***	0.917
Training	School_Board	0.433	0.072	5.987	***	0.540
School_Activities	Training	0.113	0.033	3.392	***	0.292
Communication	Home	0.127	0.023	5.433	***	0.646
Communication	Belonging	0.108	0.020	5.313	***	0.607
Belonging	Home	0.149	0.027	5.603	***	0.653
e8	e7	0.305	0.065	4.700	***	0.386
e9	e7	0.222	0.066	3.351	***	0.254
e13	e12	0.281	0.075	3.749	***	0.321
e36	e39	0.193	0.051	3.760	***	0.291
e42	e44	0.138	0.031	4.416	***	0.384
e51	e52	0.139	0.028	4.968	***	0.382
e58	e59	0.256	0.065	3.931	***	0.289
e59	e60	0.144	0.049	2.928	0.003	0.203
e60	e62	0.332	0.063	5.257	***	0.475
e14	e7	0.279	0.084	3.324	***	0.226
e5	e19	0.188	0.052	3.574	***	0.277
e15	e13	0.156	0.069	2.246	0.025	0.168
e28	e29	0.155	0.025	6.138	***	0.557
e58	e45	0.121	0.055	2.182	0.029	0.133
e58	e67	0.189	0.038	4.929	***	0.395
e64	e71	0.139	0.043	3.249	0.001	0.273
e45	e51	0.135	0.038	3.530	***	0.236
e53	e52	0.120	0.028	4.249	***	0.310
e49	e53	0.083	0.036	2.313	0.021	0.169
e45	e46	0.167	0.054	3.116	0.002	0.233

**** Statistical significance of the regression between each item and the assigned construct.*

barely been developed in Spain, justifying the diversity of covariance between measurement errors. While neighborhood activities (Q60) may be ecological (Q59) or religious (Q62), according to Yulianti et al. (2019), these have been differentiated between, since they can be carried out by other associations. On the other hand, collection activities (Q58) are civic and solidarity-based, being external to the school environment, whereas textbook collection and distribution (Q45) is internal, and ecological activities (Q59) do not pursue collection, but rather, active involvement. All of these "are important with a strong range of institutions (religious, recreational, corporate, and library) that are linked to maintain stability, cohesion, and well-being of the community" (Gahwaji, 2019, p. 11).

TABLE 6 | Goodness of fit indices of the IMFIS-TP.

Index	Desired value	Obtained value
CMIN/DF	Between 1 and 5	1.88
CFI	≥0.9	0.90
RMSEA	<0.08	0.07

Finally, "promoting family training" has covariance of measurement errors with variables from the previous subscale (Q64–Q71 and Q58–Q67), since according to IMFIS-TP, they are related to one another. In fact, taking advantage of free time (Q64) is one of the recurrent topics of the training offer (Q71) (Hernández-Prados and Álvarez-Muñoz, 2019).

TABLE 7 | Overall reliability of the questionnaire and reliability by dimensions.

Dimensions	Cronbach's α	McDonald's Ω
Overall	0.975	0.982
Communication with the school	0.731	0.703
Participation in school activities	0.875	0.840
Sense of belonging	0.915	0.899
Involvement in the home	0.921	0.904
Involvement in the PA and School Board	0.975	0.959
Community Participation	0.874	0.889
Training	0.916	0.816

The creation of this instrument can allow us to know the reality of the teacher and promote measures for training from the university (initial education), teacher training centers, and internal training promoted by the school (permanent training).

Based on all of this, the study has demonstrated that the techniques developed to assess the structural equations model have a confirmatory bias. Therefore, although the proposed model has an acceptable fit, the researcher has not tested it, but rather, has only confirmed that it is one of the various potential models (Cupani, 2012). As for the study's limitations, although the QFIS-TP suitably fits the IMFIS-TP and permits knowledge of the level of competency of the teachers as promoters of family participation, future works should expand and diversify the sample of teachers, both in a national and international scope. Longitudinal studies are needed, in contrast with the qualitative studies, to identify other actions that teachers promote and that may not have been considered in the QFIS-TP. All of this will favor the creation of instruments that offer knowledge having a better fit with the teacher's reality and will promote training measures from the university (initial education), the teacher training centers, and internal training promoted by the school. It would also be interesting to determine the teacher's level of belonging to the school, since they cannot promote this if they themselves do not experience it. All of this suggests that this emerging field demands further research to help improve family participation.

DATA AVAILABILITY STATEMENT

The data are available and will soon appear in the Centro de Investigaciones Sociológicas (CIS) database called ARChivo de Estudios Sociales (ARCES) which can be accessed through http://www.arces.cis.es/arces.jsp.

AUTHOR CONTRIBUTIONS

MGO: exhaustive review of theoretical contributions, active collaboration in the construction of the questionnaire, general and specific proposals regarding the methodology, writing of the discussion section, and conclusions based on the results obtained. JP and MG-S: methodological development and analysis of results. MH-P: review for the theoretical basis, discussion, and conclusions. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the MINECO (Ministerio de Economía, Industria y Competitividad) under Grant number EDU2016-77035-R and the under Grant number PID2020-113505RB-I00 funded by MCIN/AEI/ 10.13039/501100011033 (Ministerio de Ciencia e Innovación/ Agencia Estatal de Investigación).

REFERENCES

Anderson, K. J., and Minke, K. M. (2007). Parent involvement in education: toward an understanding of parents' decision making. *J. Educ. Res.* 100, 311–323. doi: 10.3200/JOER.100.5.311-323

Aguinis, H., Gottfredson, R. K., and Joo, H. (2013). Best-practice recommendations for definind, identifying, and handling outliers. *Organ. Res. Methods* 16, 270–301. doi: 10.1177/1094428112470848

Benner, A. D., Boyle, A. E., and Sadler, S. (2016). Parental involvement andadolescents' educational success: The roles of prior achievement andsocioeconomic status. *J. Youth Adolesc.* 45, 1053–1064. doi: 10.1007/s10964-016-0431-4

Caldas, S. J., and Cornigans, L. (2015). Race/ethnicity and social capital among middle-and upper-middle-class elementary school families: a structural equation model. *Sch. Comm. J.* 25, 137–156.

Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E., and Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educ. Res. Rev.* 14, 33–46. doi: 10.1016/j.edurev.2015.01.002

Cheung, C. K., Cheung, H. Y., and Hue, M. T. (2017). Educational contributions to students belongingness to the society, neighbourhood, school and family. *Internat. J. Adolesc. Youth* 22, 226–237. doi: 10.1080/02673843.2016.1157827

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. New York, NY: Academic Press.

Consejo Escolar del Estado (2014). *La participación de las familias en la educación escolar*. Madrid: Subdirección General de Documentación y Publicaciones.

Cupani, M. (2012). Análisis de ecuaciones estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación. *Revista Tesis* 1, 186–199.

Curran, P. J., West, S. G., and Finch, J. F. (1996). The robustness of test statistics to nonnormal- ity and specification error in confirmatory factor analysis. *Psychol. Methods* 1, 16–29. doi: 10.1037/1082-989X.1.1.16

Dettmers, S., Yotyodying, S., and Jonkmann, K. (2019). Antecedents and outcomes of parental homework involvement: how do family-school partnerships affect parental homework involvement and student outcomes? *Front. Psychol.* 10:1048. doi: 10.3389/fpsyg.2019.01048

DeVellis, R. F. (2003). *Scale development: Theory and applications*. 2nd ed. Thousand Oaks, CA: Sage.

Ebbert, A. M., Infurna, F. J., and Luthar, S. S. (2019). Mapping developmental changes in perceived parent–adolescent relationship quality throughout middle school and high school. *Dev. Psychopathol.* 31, 1541–1556. doi: 10.1017/S0954579418001219

Epstein, J. L. (2011). *School, Family and Community Partnerships. Preparing Educators and Improving Schools*. Boulder: WESTVIEW Press.

Epstein, J. L., Sanders, M. G., Sheldon, S. B., Simon, B. S., Salinas, K. C., Jansorn, N. R., et al. (2019). *School, family, and community partnerships: Your handbook for action*. Thousand Oaks, CA: Corwin press.

Erdener, M. A. (2016). Principals' and Teachers' Practices about Parent Involvement in Schooling. *Univ. J. Educ. Res.* 4, 151–159. doi: 10.13189/ujer.2016.041319

Europe Strategy 2020 (2015). *Unión por la innovación*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas.

European Commission (2000). *Informe europeo sobre la calidad de la educación escolar. Dieciséis indicadores de calidad*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas.

European Commission (2005). *Propuesta de Decisión del Parlamento Europeo y del Consejo de Europa que dé luz verde al programa "Ciudadanos para Europa" durante el periodo 2007-2013 para promover una ciudadanía europea activa*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas.

Gahwaji, N. M. (2019). The implementation of the epstein's model as a partnership framework at saudi kindergartens. *J. Coll. Teach. Learn.* 16, 11–20. doi: 10.19030/tlc.v16i2.10318

García-Sanz, M. P., Hernández-Prados, M. A., Galián-Nicolás, B., and Belmonte-Almagro, M. L. (2020). Docentes, familias y órganos de representación escolar. *Estudios Educ.* 38, 125–144.

Garbacz, S. A., Hall, G. J., Young, K., Lee, Y., Youngblom, R. K., and Houlihan, D. D. (2019). Validation study of the family involvement questionnaire–elementary version with families in belize. *Assess. Effect. Interv.* 45, 73–80. doi: 10.1177/1534508419862857

Gomariz, M. A., Hernández-Prados, M. A., García-Sanz, M. P., and Parra, J. (2017). Tejiendo puentes entre la escuela y la familia. El papel del docente. *Bordón Rev. Pedagog.* 68, 41–57. doi: 10.13042/Bordon.2016.49832

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (2008). *Análisis multivariante*. 5th ed. Madrid: Prentice Hall.

Hernández-Prados, M. A. (2019). "El encuentro formativo de las familias y docentes. Un espacio de colaboración," in *Formar para Transformar. Cambio Social y Profesiones Educativas*, ed. J. Vera-Vila (Granada: GEU), 267–270.

Hernández-Prados, M. A., and Álvarez-Muñoz, J. S. (2019). Family leisure and academic achievement. Perception of the families. *Ital. J. Educ. Res.* 86–105. Available online at: https://ojs.pensamultimedia.it/index.php/sird/article/view/3685

Hernández-Prados, M. A., García-Sanz, M. P., Galián-Nicolás, B., and Belmonte-Almagro, M. L. (2019). Implicación de familias y docentes en la formación familiar. *Rev. Electrón. Interuniv. Formación Profesorado* 22, 61–75. doi: 10.6018/reifop.388971

Hornby, G. (2011). *Parental involvement in childhood education: Building effective school-family partnerships*. New York, NY: Springer Science & Business Media. doi: 10.1007/978-1-4419-8379-4

Hu, L. T., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psycholog. Methods* 3, 424–453. doi: 10.1037/1082-989X.3.4.424

Kline, R. B. (2005). *Principies and practice of structural equation modeiing.* 2nd ed. New York, NY: Guilford.

Kurtulmus, Z. (2016). Analyzing parental involvement dimensions in early hildhood education. *Educ. Res. Rev.* 11, 1149–1153. doi: 10.5897/ERR2016.2757

Ladky, M., and Peterson, S. S. (2008). Successful practices for immigrant parent involvement: An Ontario perspective. *Multicult. Perspect.* 10, 82–89. doi: 10.1080/15210960801997932

Landis, R., Edwards, B. D., and Cortina, J. (2009). "Correlated residuals among items in the estimation of measurement models," in *Statistical and methodological myths and urban legends: Doctrine, verity, and fable in the organizational and social sciences*, eds C. E. Lance and R. J. Vandenberg (Milton Park: Routledge), 195–214.

Lévy, J. P., and Varela, J. (2003). *Análisis multivariante para Ciencias Sociales*. Madrid: Pearson-Prentice Hall.

Lingard, B., Baroutsis, A., and Sellar, S. (2014). *Learning Commission Report: Connecting schools with communities*. Brisbane: The University of Queensland.

Maxwell, B., and Staring, F. (2018). *Better learning for Europe's young people: developing coherent quality assurance strategies for school education*. Bruselas: Comisión Europea.

Organisation for Economic Co-operation and Development [OECD] (2011). When students repet grades or are transferred out of school: what does it mean for education systems? *PISA Focus* 6, 1–4.

Organisation for Economic Co-operation and Development [OECD] (2012). *Let's Read Them a Story! The Parent Factor in Education*. Paris: PISA, OECD Publishing.

Organisation for Economic Co-operation and Development [OECD] (2014). *Panorama de la educación 2014*. Paris: PISA, OECD Publishing.

Severiens, S., Wolff, R., and van Herpen, S. (2014). Teaching for diversity: a literature overview and an analysis of the curriculum of a teacher training college. *Eur. J. Teach. Educ.* 37, 295–311. doi: 10.1080/02619768.2013.845166

Spanish Strategy of Science and Technology and of Innovation [EECTI] (2012). https://www.ciencia.gob.es/Estrategias-y-Planes/Estrategias/Estrategia-Espanola-de-Ciencia-Tecnologia-e-Innovacion-2021-2027.html;jsessionid=D52CD16D241D56361C80ED233EAF4A8D.2

Tran, Y. (2014). Addressing reciprocity between families and schools: Why these bridges are instrumental for students' academic success. *Improv. Sch.* 17, 18–29. doi: 10.1177/1365480213515296

Uslu, F., and Gizir, S. (2017). School belonging of adolescents: The role of teacher–student relationships, peer relationships and family involvement. *Educ. Sci.* 17, 63–82. doi: 10.12738/estp.2017.1.0104

Verdugo, M. A., Crespo, M., Badía, M., and Arias, B. (2008). *Metodología en la investigación sobre discapacidad. Introducción al uso de las ecuaciones estructurales.* Paper presented at the VI Seminario Científico SAID, Salamanca, june, 5 and 6.

Vogels, R. (2002). *Ouders bij de Les. Betrokkenheid van Ouders bij de School van hun Kind*. Den Haag: Sociaal Cultureel Planbureau.

Walker, J. M. T., Wilkins, A. S., Dallaire, J. R., Sandler, H. M., and Hoover-Dempsey, K. V. (2005). Parental involvement: Model revision through scale development. *Elem. Sch. J.* 106, 85–105. doi: 10.1086/499193

Yulianti, K., Denessen, E., Droop, M., and Veerman, G. (2019). Transformational Leadership for Parental Involvement: How Teachers Perceive School Leadership Practices to Promote Parental Involvement in Children's Education". *Leadership Policy Sch.* 18, 1–16. doi: 10.1080/15700763.2019.1668424

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Education

ORIGINAL RESEARCH
published: 08 March 2022
doi: 10.3389/feduc.2022.815795

Impact of School Population Composition, Workload, and Teachers' Utility Values on Teaching Quality: Insights From the Dutch TALIS-2018 Data

Kim H. R. Ouwehand[1], Kate M. Xu[2], Marieke Meeuwisse[1], Sabine E. Severiens[1] and Lisette Wijnia[2]*

[1] Department of Psychology, Education and Child Studies, Erasmus School of Social and Behavioral Sciences, Erasmus University Rotterdam, Rotterdam, Netherlands, [2] Department of Conditions for Lifelong Learning, Faculty of Educational Sciences, Open Universiteit Nederland, Heerlen, Netherlands

OPEN ACCESS

Edited by:
James Ko, The Education University of Hong Kong, Hong Kong SAR, China

Reviewed by:
Christoph Helm, University of Teacher Education Zug, Switzerland
Karien Dekker, RMIT University, Australia

***Correspondence:**
Kim H. R. Ouwehand
ouwehand@essb.eur.nl

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Education

Received: *15 November 2021*
Accepted: *02 February 2022*
Published: *08 March 2022*

Citation:
Ouwehand KHR, Xu KM, Meeuwisse M, Severiens SE and Wijnia L (2022) Impact of School Population Composition, Workload, and Teachers' Utility Values on Teaching Quality: Insights From the Dutch TALIS-2018 Data. Front. Educ. 7:815795. doi: 10.3389/feduc.2022.815795

Several studies show that teaching quality is an important predictor of students' academic achievement. However, less is known about factors that are important for teaching quality. In the present study, it was hypothesized that school population composition [i.e., students' socioeconomic status (SES) and migration background], workload, and teachers' utility values toward teaching would be important factors related to their teaching quality. The Dutch *Teaching and Learning International Survey* (TALIS) 2018 was explored (N = 1,884, secondary school teachers and 116 school leaders) to test our hypotheses. Data gathering followed a two-step procedure. Firstly, 200 schools were randomly selected. Secondly, 20 teachers within each school were randomly selected. Multi-item scales operationalized teaching quality on teachers' self-reported classroom management, cognitive activation, clarity of instruction, and positive student-teacher relationships. Multilevel analyses showed that teachers' social utility value was positively associated with all dimensions of teaching quality, whereas personal utility value was only associated with classroom management and clarity of instruction. Teachers working at schools with moderate shares of students from a socioeconomically disadvantaged background reported higher clarity of instruction (β = 0.42), and moderate and high shares report lower cognitive activation (β = –0.40, β = –0.33, respectively) than those working at schools with low shares. Student-teacher relationships were rated more positively by teachers working at schools that reported no students with a migrant background than those working at schools with a small share (β = 0.33). Moreover, teachers working at schools with high shares of students with migration backgrounds (β = –0.17) reported more negative relationships. These results suggest that dealing with low-SES students at schools affects the cognitively-focused elements of teaching quality, while dealing with students with a migration background seems to affect the social teaching qualities of teachers.

Keywords: SES, migration background, classroom management, cognitive activation, clarity of instruction, student-teacher relationship

INTRODUCTION

High-quality education is crucial in our modern society. School effectiveness, often determined by students' performance attainments at the end of schooling, can be studied using the Context-Input-Process-Outcome (CIPO) model (Scheerens, 1990, 2000). According to this framework, the output of education, such as students' achievement, cannot be studied in isolation from input and process variables within a particular context. Input variables can consist of student, classroom, or teacher characteristics and financial and material aids in a school (Scheerens, 2000; Klieme, 2013). The transformation process between input variables and the output of schooling consists of the factors that make it possible for students to learn, such as instructional methods and curriculum decisions. Because teachers are at the center of the learning materials and the students, they are key figures in promoting students' motivation and achievement (e.g., Wentzel, 2009). Indeed, research shows that teaching quality is crucial for students' academic achievement (e.g., Baumert et al., 2010; Higgins et al., 2014; Gustafsson et al., 2018) and motivation (Maulana et al., 2016; Aelterman et al., 2019). Therefore, teaching quality can be considered a critical process variable within the CIPO model (Klieme, 2013).

Teaching quality is a multidimensional construct (Lazarides et al., 2021). Klieme et al. (2009) and Fauth et al. (2019) identified three dimensions of teaching quality: classroom management, cognitive activation, and supportive climate. Classroom management involves coping with disruptive behaviors, expressing and implementing classroom rules and procedures, and smooth transitions. Effective classroom management is considered a precondition for effective teaching and learning (Jones and Jones, 2012). Classroom management interventions have been associated with various adaptive student outcomes, such as academic achievement and time-on-task (Korpershoek et al., 2016). Cognitive activation concerns the use of challenging tasks and practices that aim to foster students' cognitive engagement, such as discussions and student activation, and has been positively associated with academic achievement (Fauth et al., 2019). The third teaching quality dimension, supportive climate, consists of a student-oriented affective and a cognitive component (Klieme et al., 2009; Lazarides et al., 2021). The affective component concerns positive student-teacher relations that are warm, caring, and emotionally supportive (Fauth et al., 2019; Lazarides et al., 2021). The cognitive component concerns the clarity of instruction. It refers to a teacher's ability to explain content clearly to students, such as setting clear goals and expectations and making connections with prior knowledge (Organisation for Economic Co-Operation and Development [OECD], 2019; Lazarides et al., 2021). Fauth et al. (2019) showed that the teaching quality dimension supportive climate was positively associated with students' interest.

Because teaching quality dimensions are conditional for learning and have been associated with positive student outcomes, it is crucial to investigate which school and teacher factors affect teaching quality. Therefore, in this study, we examined the effect of teachers' utility values, workload, and school population composition on teaching quality. Teaching quality was defined by the four components described above: classroom management, cognitive activation, clarity of instruction (i.e., the cognitive component of supportive climate), and positive student-teacher relations (i.e., the affective dimension of supportive climate; Klieme et al., 2009; Fauth et al., 2019; Lazarides et al., 2021). All these components are assessed by the Dutch Teaching and Learning International Survey (TALIS) 2018 dataset, consisting of 1884 secondary school teachers and 116 schools. TALIS is a reoccurring international survey organized by the OECD and has been set out in 2008, 2013, and 2018. The next round is planned for 2024. For TALIS 2018, 260,000 teachers and 15,000 school leaders participated, representing 48 countries and economies. In separate surveys, teachers and school leaders rate their perceptions on work-related constructs such as professional development, teaching beliefs, practices, work assessment, job satisfaction, needs, workload, leadership, work climate, and school population composition (Organisation for Economic Co-Operation and Development [OECD], 2019).

School Population Composition

School population composition is considered as an input variable in our model. The school population composition, such as the share of students with a low socioeconomic status (SES) and migration backgrounds of its students, could impact teaching quality. For the TALIS, migration background was defined by being born in another country or having parents born in another country. Important to mention is that SES and migration background overlap as immigrants or people with a (non-western) migration background have relatively low SES when compared to people without a (non-western) migration background (Geerlings et al., 2018). However, studies show that effects of SES and migration background can be disentangled for some teaching quality characteristics. For example, Rjosk et al. (2014) showed that teachers tend to decrease cognitive activation (i.e., challenging students' thinking) when teaching in classrooms with a high number of students coming from low-SES backgrounds (regardless of whether they were from a language minority background). It was also found that teachers perceive high SES students as more teachable than low-SES students, even when cognitive abilities are equal (Agirdag, 2018). Furthermore, schools with a relatively high amount of low-SES students also spend less instruction time on challenging domains, such as science (Willms, 2010). These studies suggest that a higher share of low-SES students at a school is related to lower quality and quantity of instruction.

In the Netherlands, the cultural diversity of the population has increased massively; the population with an immigrant status or migrant background has grown from 9.2% in 1972 (Jennissen et al., 2018) to 24.8% in 2020 (Centraal Bureau voor de Statistiek [CBS], 2020). More culturally diverse schools and classrooms have emerged with this increase, which can be challenging for Dutch teachers (for an overview of the Dutch educational system,

see Nuffic, n.d.).[1] In a multicultural classroom, teachers need to be aware of the different factors that influence classroom life and need to address the needs and interests of students with different backgrounds (Wubbels et al., 2006; Patall and Zambrano, 2019). Therefore, a multicultural classroom requests higher levels of teaching quality than homogenous classrooms (Wubbels et al., 2006; Severiens et al., 2014; Archambault et al., 2020). Wubbels et al. (2006) identified four competencies teachers need in multicultural classrooms: (a) Monitoring and managing student behavior, (b) creating positive student-teacher relationships, (c) teaching for student attention and engagement, and (d) required teacher attitudes and knowledge. These competencies align with generic teaching quality dimensions (Wubbels et al., 2006; Severiens et al., 2014). For example, monitoring and managing student behavior is similar to classroom management (Klieme et al., 2009). The competence creating positive student-teacher relationships concerns monitoring the needs of their students, showing respect for and interest in the students and their background (Wubbels et al., 2006), and aligns with the cognitive and affective dimensions of supportive climate (Lazarides et al., 2021). The third competence, teaching for student engagement and attention, concerns stimulating all students to participate and relies on cognitive activation (Severiens et al., 2014). The fourth component, having the required attitudes and knowledge, means that teachers need to be aware of their own normative behavior, their perceptions of students, and the position of students within their class (Wubbels et al., 2006). This last component is not included in the conceptualizations of teaching quality by Klieme et al. (2009). Drawing further on this fourth competence, Tielman et al. (2021) stressed the tension field between the normative values of teachers and that of pupils in culturally diverse classes "Teachers sometimes find themselves in a struggle between their own values and changing value systems in society on the one hand, and at the same time their wish to let students develop their own value orientations" (p. 3). Tielman et al. (2021) concluded that student-teacher relationships based on trust and a pleasant social climate could decrease these tensions between teachers and students. This study underscores the importance of establishing a supportive climate with positive student-teacher relationships. However, in this respect, it is problematic that the share of teachers with a migration background in the Netherlands is small relative to the number of students with a migration background, especially those with a non-western background (Centraal Bureau voor de Statistiek [CBS], 2020). Thijs et al. (2012) showed that ethnic-Dutch teachers evaluated their relationship quality with Moroccan-Dutch students as less favorable than with ethnic-Dutch and Turkish-Dutch students. This suggests that ethnic incongruence between the teacher and student could affect the student-teacher relationship. Unfortunately, the TALIS did not ask teachers about their own migration or ethnic background, so no hypotheses on this could be tested in the current study.

From the research findings described above, it is expected that low SES and migration background might have similar effects but also unique effects on teaching quality. Because prior research revealed that in low-SES schools, less cognitive activation took place (Rjosk et al., 2014) and less time was spent on challenging topics (Willms, 2010), we, in particular, expected a negative effect of low SES on cognitive activation (H1). In addition, low-SES students are perceived as less teachable (Agirdag, 2018). This might lead teachers to try and compensate for their students' perceived lower cognitive abilities by increasing the clarity of instruction. Therefore, we expected higher clarity of instruction for low-SES schools (H2). From the negative impact of ethnic discrepancy between teacher and students (Thijs et al., 2012) with a dominantly white teacher population in the Netherlands (Traag, 2018) we expect more negative ratings on the student-teacher relationships for teachers working in schools with high shares of students with migration backgrounds (H3). As relational aspects between the students in ethnic diverse classrooms are tense (e.g., Tielman et al., 2021), teachers in classrooms with a higher share of students with a migration background pay more attention to classroom management (H4).

Workload

Another input variable in our model that could affect teaching quality is the experienced workload (Dube-Xaba and Makae, 2022). Workload concerns all activities teachers need to spend on their official duties during or after school hours (Johari et al., 2018). High psychological demands (Harmsen et al., 2019) and workload (Den Brok et al., 2017) are significant predictors for the attrition of teachers in the Netherlands, a country facing a serious teacher shortage already (Vereninging van Scholen in het Voortgezet Onderwijs [VO RAAD], n.d.). A study from the neighboring country Flanders, Belgium, by Amitai and Van Houtte (2022) showed that novice teachers' workload and feeling unprepared to manage diverse classrooms were reasons to leave the profession of teaching. However, well-performing teachers are less likely to quit their jobs (Krieg, 2006). Also, the supposed negative impact of workload on educational quality has been a popular item in the Dutch press (e.g., Chaudron and Dujardin, 2019). Workload is negatively associated with teachers' self-ratings of their performance and job satisfaction (Huyghebaert et al., 2018), which could, in turn, negatively affect their teaching quality. Therefore, it was hypothesized that workload negatively affects all teaching quality outcome variables (H5).

Motivation: Utility Values

A teacher's motivation to choose teaching as a career is another input variable that could affect teaching quality. Personal and social utility values were introduced by Watt and Richardson (2007) to investigate teachers' motivations to enter the profession in a system called the Factors Influencing Teaching Choice (FIT-Choice). The FIT-Choice model is based on expectancy-value

[1] In order to be able to contextualize the situation in the Netherlands, it is important to consider two specific aspects of the Dutch education system. The Dutch school system is a public school system, almost all Dutch schools are funded and monitored by the government. At the age of 12, Dutch pupils transition to one of three tracks in secondary education. Test scores in combination with a school advice, leads to either the vocational track (4 years), a general track (5 years) or a pre-university track (6 years; see also Nuffic, n.d.). Parents' SES play a large role in this advice, which impacts school composition in secondary education: the vocational track generally contains larger percentages of students from low-SES and migration backgrounds.

theory. A distinction is made between primary motivational beliefs for teaching (e.g., abilities to teach effectively) and three subjective task values: intrinsic value, personal utility value, and social utility value (Richardson and Watt, 2014). Intrinsic value refers to a teachers' interest in teaching. Personal utility value concerns the motivation for practical and attractive aspects of teaching, such as having a steady job with fixed hours and a pension plan that fits well with the teachers' personal life (e.g., time for family; Watt and Richardson, 2007). In contrast, social utility value represents more altruistic motivation, such as contributing to society, shaping young students' future, and enhancing social equity (Watt and Richardson, 2007, 2012; Richardson and Watt, 2014).

In our study, we were interested in the effect of teachers' subjective task values on teaching quality. In the TALIS, only personal and social utility values were measured (Organisation for Economic Co-Operation and Development [OECD], 2019). Social utility value is considered an adaptive motivation for teaching (Richardson and Watt, 2014). Research shows that social utility values have been associated with pre-service teachers' interest in the profession (Giersch, 2016), planned persistence to become a teacher, and retention (O'Brien and Schillaci, 2002; Bruinsma and Jansen, 2010; Richardson and Watt, 2014). There is some evidence that social utility value is positively associated with teaching quality. Research showed that social utility value was directly associated with better student-teacher relationships and certain aspects of classroom management (Berger and Girardet, 2021). Moreover, Richardson and Watt (2014) found that social utility value predicted later positive teaching behaviors (e.g., positive expectations, student-teacher relationships, and clear structure/expectations) through planned persistence. Furthermore, with qualitative interviews, Parr et al. (2021) found links between social utility value and favorable instructional practices such as challenging students to think critically and centering instruction on students' needs. Therefore, we expected that social utility value is positively associated with all teaching quality dimensions (H6).

In contrast to social utility values, research showed more mixed findings for personal utility values. For example, König and Rothland (2012) found a negative association between "job security" motives with pedagogical knowledge of pre-service teachers on a first measurement, but a positive effect on learning gain over the course of 1 year. In contrast, Girardet and Berger (2018) showed that teachers with personal utility values showed a less adaptive change in classroom management style focusing on structure after a teacher education program. Positive associations of personal utility values have been found with career interest (Giersch, 2016) and job satisfaction (Liu et al., 2020). However, Richardson and Watt (2014) found that personal utility value was not associated with planned persistence and positive and negative teaching behaviors. Due to the mixed effects of prior research, we expected that personal utility would either be unassociated with teaching quality dimensions or would show small, positive associations (H7). However, we did not expect a negative effect on teaching quality.

As mentioned, personal utility values concern the practical and attractive aspects of teaching as a reason for entering a teaching career, such as work-life balance and job security (Watt and Richardson, 2007). A higher workload could impact the actual benefits a teacher reaps from having a teaching career. Increased workload means the job could affect the time available for private life (cf. Johari et al., 2018). It was hypothesized that workload moderates the relationship between personal utility value on teaching quality. Specifically, we expected that high workload combined with high personal utility value could negatively affect teaching quality (H8). No such moderating effect of workload is expected on the association between social utility value and teaching quality (9).

Cross-Level Interactions

The hypothesized relationships between workload and utility values, on the one hand, and teaching quality, on the other hand, may differ in schools that vary according to composition. In other words, we may find cross-level interactions. The reason is that schools with large shares of students from low SES and migration backgrounds more often deal with problems such as high rates of teacher dropout in combination with a relatively high number of social problems among their students (Payne, 2014). Moreover, teacher self-efficacy is often lower in these schools (Parkhouse et al., 2019). Teachers in these schools may have high personal and social utility values. However, in a context of low teacher retention combined with low levels of teacher self-efficacy, teachers' high utility values may not translate as easily into an orderly classroom climate, clear instruction, and good quality student-teacher relationships.

It is also not unthinkable that a relatively high workload in classrooms with high shares of students from low SES and migration backgrounds makes it even harder to establish good teaching quality. Or conversely, in urban schools that manage to keep workload relatively low, it may be possible to deliver good quality instruction, good classroom management, student-teacher relationships, and cognitive activation. Based on these lines of reasoning, we arrive at two hypotheses with regard to cross-level interactions. Firstly, in classrooms with relatively low shares of low-SES students and (or) students from migration backgrounds, we expected utility values to affect teaching quality positively. In contrast, this relation might be less pronounced in classrooms with relatively high percentages of low-SES students and (or) students from migration backgrounds (H10). Secondly, we expected a negative relationship between workload and teaching quality in classrooms with relatively high percentages of low-SES students and (or) students from migration backgrounds. In contrast, this relationship could be less pronounced in classrooms with relatively low shares of low SES-students and (or) students from migration backgrounds (Hypothesis 11).

Hypotheses of the Present Study

Below we provide an overview of all hypotheses specified and argued for in the introduction:

H1: Lower cognitive activation at low-SES schools.
H2: Higher clarity of instruction at low-SES schools.

H3: Lower quality of student-teacher relationships at schools with high shares of students with a migration background.
H4: Higher classroom management skills at schools with high shares of students with a migration background.
H5: Negative effect of workload on all teaching quality outcome variables.
H6: Positive effect of social utility and all teaching qualities.
H7: No (or small) effects of personal utility values on teaching quality.
H8: Workload moderates relation between personal utility and teaching quality.
H9: No moderation effect of workload on the relation between social utility and teaching quality.
H10: In classrooms with relatively low shares of low-SES students and (or) students from migration backgrounds, both utility values have a positive effect on teaching quality.
H11: In classrooms with relatively high shares of low-SES students and (or) students from migration backgrounds, there is a negative relation between workload and teaching quality.

MATERIALS AND METHODS

Participants

In total, 1884 secondary school teachers (54% male) and 125 school leaders (62% male) from 116 schools participated in the Dutch TALIS 2018 (Sapulete et al., 2018). Teachers' age ranged between 19 and 71 years, with M_{age} = 42.8 years, SD_{age} = 11.85. School leaders' age ranged between 32 and 68 years, with M_{age} = 53.9 years, SD = 7.76. Response rates were 76% among teachers and 80% among school leaders.

Procedure

Data gathering followed a two-step procedure. Firstly, 200 schools were randomly selected. Secondly, 20 teachers within each school were randomly selected (see TALIS 2018, Technical Report for a detailed description; Organisation for Economic Co-Operation and Development [OECD], 2019). To be included in the analysis of TALIS 2018, at least 50% of teachers within each selected school had to participate. The teacher and school leader surveys were administered between March and July 2018 (for more information, see Organisation for Economic Co-Operation and Development [OECD], 2019).

Variables

Below, we describe the (sub)scales used in this study and their abbreviated names in the dataset (see TALIS 2018 technical report; Organisation for Economic Co-Operation and Development [OECD], 2019). Reliability is indicated by Cronbachs alpha (α) for sets of items with tau-equivalence (all the items have equal covariance with the true score), and omega coefficients when tau-equivalence was violated (see Deng and Chan, 2017).

Level 2 (School Level) Independent Variables

The shares of students from a socioeconomically disadvantaged background (SES) and students with an immigrant status or migration background (MIGR) were based on the principals' ratings on a five-point Likert scale (1 = none, 2 = 1–10%, 3 = 11–30%, and 4 = 31–60%, or 5 $\geq$ 60%). With two separate items, principals were asked to estimate the share of (1) students from socioeconomically disadvantaged families and (2) those with an immigrant status or migration background. They received short defining descriptions on both terms. Socioeconomically disadvantaged families were described as unprivileged families that need help (social services, help organizations) to fulfill their basic needs, such as shelter, food and health care. An immigrant student was described as a student that is born abroad and students with a migration background as having parents who are born abroad.

The ratings on SES and migration background were converted into dummies, with the largest category being the reference group. The majority of schools reported having 1–10% students of low-SES backgrounds and 1–10% students with a migration background. Therefore, for both SES and MIGR, the reference group was 1–10% students. For SES, two dummies were created. The first dummy, SES1, contrasted the reference group with schools with 11–30% low-SES students, and the second dummy, SES2, contrasted the reference group with schools with 31–60% and > 60% low-SES students. These two final categories were piled together because the > 60% category contained only one school with 11 teachers. No schools were reported to be without students from low-SES backgrounds.

For migration background, three dummies were created. Nine schools reported having no students with a migration background. MIGR0 contrasted the reference group (1–10% migration background) with schools without immigrant status or migration background students. MIGR1 contrasted the reference group with schools reporting 11–30% of students with immigrant status or migration backgrounds. Finally, MIGR2 contrasted the reference group with schools reporting 31–60% and > 60% of students with immigrant status or migration background. These two final categories were piled together because the > 60% category contained only one school with 11 teachers (see **Table 1** for the frequencies of SES and Migration per category).

Level 1 (Teacher Level) Independent Variables

Social utility value was composed of three items (α = 0.78), personal utility value by four items (α = 0.93), and workload by five items (α = 0.74). In the analysis, social utility value, personal utility value, and workload were centered around the grand mean (see **Appendix** for all the items).

Dependent Variables: Teaching Quality

Classroom management (α = 0.89), cognitive activation (α = 0.79), clarity of instruction (α = 0.74), and student-teacher

TABLE 1 | Distribution of schools reporting low to high shares of students with socioeconomically disadvantaged (SES) backgrounds and (or) immigrant status or migration backgrounds (MIGR).

	SES				MIGR			
	Schools		Teachers		Schools		Teachers	
Category	Number	%	Number	%	Number	%	Number	%
0%	0	0	0		9	7.1	136	7.2
1–10%	74	58.3	1,088	57.7	86	67.7	1,327	70.4
11–30%	35	27.6	559	29.7	17	13.4	261	13.9
31–60%	8	6.3	118	6.3	5	3.9	41	2.2
>60%	1	0.8	11	0.6	1	0.8	11	0.6
Missing	9	7.1	108	5.7	9	7.1	108	5.7
	127	100	1,884	100	127	100	1,884	100

relationships (Ω = 0.81) were each composed by four items (see **Appendix**).

Analyses

Firstly, it was tested whether multilevel modeling would be preferred over a unilevel approach. This was done by comparing a null model (in which we fixed variances between clusters to zero) to an independence model in which variances within and between clusters are accounted for. We used a Chi-square test of the deviance as a fit measure for our modeling, with a cut-off of Δ-2LL/df > 3.84 (which is the "95% critical value for a squared normal deviate for only one cell"; Breslow and Day, 1987, p. 130). As shown in **Table 2**, the independence model has a significantly better fit than the null model in which the variance between clusters is fixed (i.e., not accounted for). We also added the AIC and BIC to **Table 2** as model fit indices.

Secondly, we inspected the correlations between all independent and dependent variables (see **Table 3**). Two multivariate multilevel analyses using Mplus 8 were conducted. The first analyses tested the relations and possible cross-level interactions between social and personal utility value, workload, SES, and migration levels on classroom management, cognitive activation, clarity of instruction, and student-teacher relationships. It was assessed whether it would be valid to test for cross-level interactions using random slopes for the lower-level variables. This method reduces Type II errors and therefore is a strict approach to test for cross-level interactions (for an elaborate argumentation and statistical support of this method, see Heisig and Schaeffer, 2019). The second analysis tested whether workload moderated the relationship between utility values (social and personal) and classroom management, cognitive activation, clarity of instruction, and student-teacher relationships.

RESULTS

First, the results of the correlational analyses that are relevant to our hypotheses are described, followed by the multivariate multilevel regression model and moderation model testing the hypotheses. As an index for effect size, standardized β will be given for the significant results reported. All significant results regarding the correlations and hypotheses will be reported in the text for readability. All statistics can be found in **Tables 2** and **3**.

SES was positively correlated with social utility value, and clarity of instruction, indicating that teachers working at schools with more students from socioeconomically disadvantaged backgrounds reported higher levels of social utility value and spent more time on clarity of instruction. Migration background correlated positively with social and personal utility value and negatively with student-teacher relationships. This finding indicates that teachers working at schools with higher numbers of students with a migration background reported higher levels of social and personal utility value and rated the relationship with their students more negatively. Furthermore, personal utility value was positively associated with workload, classroom management, and clarity of instruction.

In support of H1, the multilevel regression model, the SES1 and SES2 contrasts yielded negative coefficients for cognitive activation, β = −0.40 and β = −0.33, respectively. This indicates that teachers at schools with a moderate and high compared to a low share of low-SES students spend less time and attention on challenging and activating teaching. In support of H2, the results revealed that the SES1 contrast was positively related to clarity of instruction, β = 0.42. This indicates that teachers at schools with a moderate share of low-SES students reported spending more time and attention on clarity of instruction than those teaching at schools with a low share of low-SES students.

In support of H3, the MIGR0 contrast showed that teachers at schools reporting to have no students with a migration background report a more positive student-teacher relationship than those working in schools with a moderate share of students with a migration background β = 0.33. The MIGR2 contrast showed that teachers working at schools with a high share of students with a migration background reported a more negative student-teacher relationship quality than those working at schools without students with a migration background, β = −0.17. In contrast to our expectations (H4), no effect on classroom management was found.

Next, the relations between the independent Level 1 (variables workload, social and personal utility values) and the outcome variables (classroom management, cognitive activation, clarity

TABLE 2 | Multivariate multilevel model testing social and personal utility values (SUV and PUV), workload (WL) and school population characteristics (SES and MIGR) on teaching quality: Clarity of instruction (CI), classroom management (CM), cognitive activation (CA) and student-teacher relationships (STR).

			0 Model	Ind model	Level 1	Level 2	β
CI	ICC = 4%	*Mean*	12.32 (0.06)***	12.32 (0.06)***	12.34 (0.05)***	12.23 (0.06)***	
Level 1	Fixed	WL			0.04 (0.03)	0.04 (0.03)	0.05 (0.03)*
		SUV			0.12 (0.02)***	0.11 (0.02)***	0.15 (0.03)***
		PUV			0.05 (0.02)*	0.05 (0.02)**	08 (0.03)**
		SUV × WL			−0.01 (0.01)	−0.02 (0.01)	−0.04 (0.03)
		PUV × WL			−0.02 (0.01)	−0.02 (0.01)	−0.04 (0.03)
	Random	σ^2_e	2.91 (0.11)***	2.79 (0.11)***	2.70 (0.11)***	2.69 (0.11)***	
		σ^2_{u0}		0.11 (0.04)**	0.09 (0.03)**	0.07 (0.04)	
Level 2		SES1				0.37 (0.14)**	0.42 (0.15)**
		SES2				0.16 (0.32)	0.10 (0.16)
		MIGR0				0.13 (0.17)	0.12 (0.15)
		MIGR1				−0.13 (0.16)	−0.11 (0.15)
		MIGR2				0.41 (51)	−0.18 (18)
CM	ICC = 1.6%	*Mean*	10.95 (0.06)***	10.95 (0.06)***	10.97 (0.06)***	11.04 (0.07)***	
Level 1	Fixed	WL			0.03 (0.03)	0.03 (0.04)	0.03 (0.03)
		SUV			0.11 (0.02)***	0.11 (0.02)***	0.13 (0.03)***
		PUV			0.06 (0.02)*	0.06 (0.02)*	0.07 (0.03)*
		SUV × WL			−0.01 (0.01)	−0.01 (0.01)	−0.03 (0.03)
		PUV × WL			< 0.01 (0.01)	−0.01 (0.01)	−0.02 (0.02)
	Random	σ^2_e	4.35 (0.18)***	4.30 (0.20)***	4.23 (0.19)***	4.18 (0.20)***	
		σ^2_{u0}		0.06 (0.05)	0.01 (0.05)	0.02 (0.06)	
Level 2		SES1				0.06 (0.15)	0.11 (0.22)
		SES2				−0.22 (0.29)	−0.10 (0.23)
		MIGR0				−0.12 (0.26)	−0.08 (0.19)
		MIGR1				−0.26 (0.15)	−0.25 (0.20)
		MIGR2				−0.52 (0.38)	−0.39 (0.24)
CA	0.7%	*Mean*	9.88 (0.06)***	9.88 (0.05)***	9.89 (0.06)***	9.99 (0.08)***	
Level 1	Fixed	WL			0.04 (0.04)	0.04 (0.04)	0.03 (0.03)
		SUV			0.011 (0.03)***	0.12 (0.03)**	0.13 (0.03)***
		PUV			0.02 (0.02)	0.02 (0.02)	0.02 (0.03)
		SUV × WL			−0.01 (0.01)	−0.01 (0.01)	−0.03 (0.03)
		PUV × WL			0.01 (0.01)	< 0.01 (0.01)	< 0.01 (0.03)
	Random	σ^2_e	4.77 (0.17)***	4.75 (0.18)***	4.69 (0.17)***	4.66 (0.18)***	
		σ^2_{u0}		0.02 (0.05)	< 0.01 (0.04)	0.02 (0.05)	
Level 2		SES1				−0.33 (0.13)*	−0.40 (0.19)*
		SES2				−0.50 (0.25)*	−0.29 (0.20)
		MIGR0				−0.20 (0.14)	−0.19 (0.19)
		MIGR1				0.07 (0.21)	0.07 (0.19)
		MIGR2				0.48 (0.47)	0.24 (0.23)
STR	ICC = 11.6%	*Mean*	13.30 (0.07)***	13.29 (0.07)***	13.30 (0.07)***	13.26 (0.10)***	
Level 1	Fixed	WL			−0.03 (0.02)	−0.03 (0.02)	−0.03 (0.03)
		SUV			0.06 (0.02)***	0.07 (0.02)***	0.10 (0.03)***
		PUV			< 0.01 (0.02)	< 0.01 (0.02)	< 0.01 (0.03)
		SUV × WL			−0.02 (0.01)*	−0.03 (0.01)*	−0.06 (0.12)
		PUV × WL			−0.01 (0.01)	−0.01 (0.01)	−0.18 (0.12)
	Random	σ^2_e	3.32 (0.11)***	2.93 (0.09)***	2.90 (0.09)***	2.93 (0.09)***	
		σ^2_{u0}		0.38 (0.09)***	0.39 (0.08)***	0.26 (0.07)***	
Level 2		SES1				−0.11 (0.15)	−0.08 (0.1)
		SES2				0.24 (0.20)	0.08 (0.11)
		MIGR0				0.80 (0.29)**	0.34 (0.10)**
		MIGR1				−0.10 (0.15)	−0.06 (0.12)
		MIGR2				−0.62 (0.24)**	−0.18 (0.12)
		-2LL (Parameters)	24705.25 (8)	24604.87 (12)	24296.79 (38)	228687.81 (64)	
		Δ-2LL(df)		100.38 (4)***	308.08 (8)***	1428.98 (26)***	
		AIC	24721.25	24628.87	24352.31	22995.81	
		BIC	24764.95	24694.42	24559.66	23341.49	

significance at the level of $p < 0.05$, **significance at the level of $p < 0.01$, and * significance at the level of $p < 0.001$.*

TABLE 3 | Correlations between the variables.

	N	MIGR	Social utility value	Personal utility value	Workload	Classroom management	Cognitive activation	Clarity of instruction	Student-teacher relationship
SES	1,776	0.48***	0.12***	0.04	−0.04	−0.01	−0.05	0.08**	−0.03
MIGR	1,776	-	0.05*	0.06*	−0.02	−0.02	0.01	0.01	−0.10***
Social utility value	1,874		-	0.05*	0.02	0.13***	0.12***	0.18***	0.07**
Personal utility value	1,873			-	0.09***	0.08**	0.03	0.08**	−0.01
Workload	1,588				-	0.04	0.04	0.05	−0.05
Classroom management	1,403					-	−0.01	0.14***	0.06*
Cognitive activation	1,403						-	0.23***	0.06*
Clarity of instruction	1,404							-	0.06*
Student-teacher relationship	1,735								-

*SES = % of students with a socioeconomically disadvantaged background; MIGR = % of students with an immigrant status or migration background; Social utility value; Personal utility value; Workload; Classroom management; Cognitive activation; Clarity of instruction; Student-teacher Relationship. *p < 0.05, **p < 0.01, ***p < 0.001.*

of instruction, and student-teacher relationship) and interaction effects (social and personal utility value x workload and personal utility value × workload) were explored. Contrary to H5, no negative relations between workload and the teaching quality variables were found. In contrast, a positive relation was found for clarity of instruction (β = 0.05). In support of H6 we found social utility value to positively relate to all teaching quality variables (for clarity of instruction, β = 0.16; classroom management, β = 0.13; cognitive activation, β = 0.13; student-teacher relationship, β = 0.10). In addition, small effects were found for personal utility value on clarity of instruction (β = 0.08) and classroom management (β = 0.07). This supports H7, stating no or small effects of personal utility value on teaching quality.

In contrast to H8, the moderation analysis showed no moderation effect on personal utility values. Unexpectedly and in contract to H9, an interaction between social utility value and workload on student-teacher relationship quality was found. Inspecting random slopes for low, medium and high levels of workload showed that the higher the workload, the more negative the relation between social utility value and student-teacher relationships.

Testing for random slopes for the lower level variables lacked support for cross-level interaction between the Level 2 (SES and MIGR) and Level 1 predictors (social and personal utility values and workload). Therefore, H10 and H11 were not supported.

DISCUSSION

Teaching quality is an important condition for learning and instruction (Jones and Jones, 2012). The present study focused on factors related to school population composition (i.e., SES and migration background), workload, and teachers' utility values (i.e., social and personal) on teaching quality (i.e., classroom management, cognitive activation, clarity of instruction, and positive student-teacher relationships). The Dutch TALIS-2018 data from secondary school teachers was used.

Firstly, in support of H1, it was found that teachers working at schools with a moderate (11–30%) or high (> 30%) share of students from a socioeconomically disadvantaged background reported spending less time and attention on a challenging and activating teaching style (i.e., cognitive activation) than their colleagues working at schools with a smaller share of low-SES students. Secondly, in line with H2 it was found that teachers working at schools with a moderate share (11–30%) of students from a socioeconomically disadvantaged background provide more clarity of instruction than their colleagues working at schools with a smaller share of low-SES students. However, this effect was not found for the high SES group. Thirdly, in line with H3, teachers working at schools with high numbers of students with an immigrant status or migration background compared to their colleagues working at schools with lower numbers, indicated to have more negative relationships with their students. In contrast to H4, no association was found with classroom management.

Interestingly, we did not find a direct negative relationship between workload and teaching quality dimensions (no support for H5). Interpreted the most straightforward way, this could mean that teachers and their teaching quality are quite robust against workload. Research suggests that workload in itself may not be a negative factor for teacher quality. Indeed, a study by Johari et al. (2018) also found no relation between workload and job performance of teachers. On the contrary, Helms-Lorenz et al. (2015) investigated interventions in the induction programs of beginning teachers and found that workload was a positive predictor of teaching quality 3 years later. In line with this, Klassen and Chiu (2010) found that teachers who experienced more workload stress reported higher classroom management efficacy than their colleagues reporting lower stress levels. It seems that the headers in the Dutch media (i.e., Chaudron and Dujardin, 2019) saying that the educational quality suffers from workload asks for a more nuanced picture. We will try to present such a picture.

The Netherlands faces an increasing teacher shortage estimated to result in 1,263 full-time equivalents in 2024 (Vereniging van Scholen in het Voortgezet Onderwijs [VO RAAD], n.d.). These numbers are problematic for Dutch secondary education. The Dutch educational inspection explains that the quality of teaching and education suffers from teacher attrition due to the increasing workload (Inspectie van het Onderwijs, 2019). A shortage of teachers can result in fewer hours in the classroom due to a lack of qualified substitutes in case of

teacher absence (e.g., sick leave) or in less quality of education due to un(der)qualified teachers in the classroom. Thus, while the present study found no relation between workload and teaching qualities in the current sample on the teacher level (all qualified teachers), educational quality, in general, is worsened by the overall teacher shortage and suboptimal solutions (unqualified teachers and canceled lessons).

In line with H6, the social utility values of teachers had a positive relationship with all teaching quality variables. In line with H7, personal utility value had a small positive effects for classroom management and clarity of instruction. Several studies indicated that social utility values are associated with teaching quality (Richardson and Watt, 2014; Berger and Girardet, 2021; Parr et al., 2021). Results for personal utility values were mixed in prior research, showing positive associations with some outcomes and unrelated associations with other outcomes. Richardson and Watt (2014) found personal utility values unrelated to positive teaching behaviors. However, König and Rothland (2012) found that pre-service teachers' personal utility values are associated with the learning gain of pedagogical knowledge. We, therefore, expected that personal utility values would be unassociated or would show (small) positive effects. Unfortunately, little research can be found looking into teacher utility values and teaching qualities (as operationalized in the present study). However we can speculate that it might be related to the need for predictability and structure of the job, underlying personal utility value. As classroom management and clarity of instruction are more qualities focusing on the structure and rules (leading to a predictable classroom situation) during teaching, teachers are motivated by predictability and structure of their job, might hang on more to these qualities. More research is needed to see if these results can be replicated in future research.

We further examined associations and potential interaction effects between workload and utility values. Correlation analyses revealed that workload was positively associated with personal utility value, indicating that teachers scoring higher on this value also indicated to experience more workload. This can be explained by the idea that the concept of a steady job with fixed hours that fits well with the teachers' personal life (e.g., holidays, part-time work; Watt and Richardson, 2007) is jeopardized by increasing workload. In contrast to H8, workload did not moderate the effect of personality utility values on teaching quality. However, we did find a unexpected (with regard to H9) interaction effect between social utility value and workload. Specifically, it was found that the relation between social utility value and student-teacher relations became more negative with increasing workload. In trying to explain this effect, we tried to look closer in what kind of workload has increased in the last decade(s). McGrath-Champ et al. (2018) point out that specifically the increased administrative load of teachers poses a serious threat to their original profession (teaching quality and student learning). We assume that especially for teachers with a high social utility value, the quality of student-teacher relationships might suffer from increasing administrative tasks.

No cross-level interactions were found which means that we found no support for H10 and H11. This indicates that the relation between social and personal utility value and workload did not differ depending on the share of low-SES students and (or) students with a migrant background.

In the present study, we examined the relationships between school composition, workload, utility values, and teaching quality using the Dutch TALIS data. These concepts are extensively theorized in international research, and insights into their relationships are interesting for an international audience. However, the Dutch educational system and the migration histories of Dutch pupils and their socioeconomic situation may lead to a specific set of results for the Netherlands. Even though our results do not seem to indicate such specificity, as most of our hypotheses can be confirmed, it is important to replicate the analyses in other national TALIS datasets to find out more about the role of specific educational systems and school population compositions.

School population composition measures for SES and migration background were at the school level and not class level. We acknowledge that shares of students with a migrant background and socioeconomically disadvantaged background at the school level may differ from the shares that teachers see in a specific classroom. TALIS does provide data on the teacher level on an estimate of SES and migration backgrounds in a specific classroom (i.e., "the first class you taught after 11 a.m. last Tuesday"). However, because all teachers teach multiple classes, these shares differ between classes. We felt that the average school-level report represent the experience of teachers on shares of low-SES and migration background students better than asking for a random class, hence the choice to look at SES and migration background at school level.

Another point for discussion is that the intraclass correlations (ICCs, see **Table 2**) are quite low for three of the four dependent variables (i.e., $< 5\%$). The decision to take a multilevel approach, despite low ICCs was based on a significantly better fit for the model when variances within *and* between clusters on the dependent variables are accounted for compared to the model in which no variance between clusters is assumed.

Moreover, the present study solely looked at the self-reports of teachers. This can be problematic because only the teacher's perspective is considered. Krammer et al. (2019) showed strong measurement invariance when comparing the teachers' ratings to that of their students. In addition, students have an advantage over the teacher when observing lessons because they are not teaching (i.e., teachers are multitasking and might be under pressure when confronted with high social density, unstructured problems, dynamic situations, and multiple demands; Scherzinger and Wettstein, 2019). Moreover, in line with Krammer et al. (2019), Scherzinger and Wettstein (2019) also showed discrepancies between teacher and student judgments of teaching quality, so the validity of only using teacher ratings can be debated. Another issue with using one method and one source of information is that found (cor)relations may be partially due to the single method (i.e., common method bias; Podsakoff et al., 2003). Future research is needed to find out whether current findings can be supported by other methods, such as looking at formal assessments of teachers' performance or classroom observations.

CONCLUSION

In conclusion, our main finding is that teachers seem to adapt their teaching quality to their teaching population. On the one hand, this is positive as students have individual needs that can be fulfilled in this manner. However, if incorrect preconceptions cause these adaptations, this can negatively affect students' outcomes (Hornstra et al., 2015; Agirdag, 2018). Indeed in the present study, we confirm earlier findings (Willms, 2010; Rjosk et al., 2014). Teachers rely less on cognitive activation and more on clarification when working at schools with a fair number of students with socioeconomically disadvantaged backgrounds. Thus, schools, where the socioeconomically disadvantaged background of students is more pronounced, might need to focus on the (pre)conceptions of the teachers about students with socioeconomically disadvantaged backgrounds and, when needed, align these (pre)conceptions correctly with the cognitive abilities of these students.

In addition, we also confirm earlier findings (e.g., Thijs et al., 2012) that ethnic diversity is a challenge for the quality of the relationship between teachers and students. More specifically, our results showed that teachers in more ethnically diverse schools reported a lower student-teacher relationship quality, reporting these relationships as less warm, caring, and emotionally supportive. Teachers' misconceptions about students' needs could partially explain this negative association. Hornstra et al. (2015) showed that teachers are less supportive of students' autonomy needs of at-risk students because they believe that students with low-SES, low ability levels, or an ethnic minority background need more teacher control. However, students' autonomy needs should be equally supported (Patall and Zambrano, 2019; Archambault et al., 2020). Therefore, more research is needed on the factors influencing their beliefs about students' SES and migration background. In addition, Wubbels et al. (2006) and Severiens et al. (2014) underscore the importance of the interpersonal competence of teachers in ethnically diverse classrooms. Our findings have implications for the focus of teacher training and professional development activities offered. Ethnically diverse schools might offer more interpersonal training focusing on diversity. Another way to improve teaching quality in ethnically diverse schools is to invest more in culturally responsive teaching (Gay, 2002). Culturally responsive teaching uses ethnically diverse students' cultural characteristics, experiences, and perspectives to teach more effectively. This way of teaching aligns well with the concept of autonomy-supportive teaching, in which teachers can take the perspective of their students (Patall and Zambrano, 2019), which is considered an aspect of good classroom management (Berger and Girardet, 2021).

In line with prior research (Richardson and Watt, 2014; Girardet and Berger, 2018; Berger and Girardet, 2021), social utility value toward teaching was related to more teaching qualities than personal utility value. This indicates that it is important to promote (pre-service) teachers' motivation to contribute to society and enhance social equity. This conclusion should be interpreted with caution as the present study limited teaching quality by the four components analyzed (i.e., classroom management, cognitive activation, clarity of instruction, and student-teacher relationships). However, we defined the chosen teaching quality components based on recurrent elements we found in our literature research described in the introduction.

In sum, the current study suggests that teachers' teaching qualities are adapted when dealing with students from socioeconomically disadvantaged backgrounds and or children with an immigrant status or migration background. Although SES and migration are correlated, different relations with teaching qualities are found in the multilevel analysis. While SES is mainly related to the cognitive aspects of teaching quality, is migration background more related to social aspects of teaching quality. Hence the main conclusion from this study is that to work on effective teaching by improving teaching quality, SES and migration need different approaches.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.oecd.org/education/talis/talis-2018-data.htm.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the OECD which was responsible for the ethical approval. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KO was part of the Dutch TALIS 2018 consortium together with MM and SS that gathered, analyzed, and reported on the Dutch TALIS 2018 data. KX was a consultant for the analyses for the study. LW was closely collaborating with KO in writing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the Dutch Ministry of Education, NRO number 504.16.385/374. The Erasmus Open Access Fonds will pay the listed article processing fee should the manuscript be accepted for publication.

REFERENCES

Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J., and Reeve, J. (2019). Toward an integrative and fine-grained insight in motivating and demotivating teaching: the merits of a circumplex approach. *J. Educ. Psychol.* 111, 497–521. doi: 10.1037/edu0000293

Agirdag, O. (2018). The impact of school SES composition on science achievement and achievement growth: mediating role of teachers' teachability culture. *Educ. Res. Eval.* 24, 264–276. doi: 10.1080/13803611.2018.1550838

Amitai, A., and Van Houtte, M. (2022). Being pushed out of the career: former teachers' reasons for leaving the profession. *Teach. Teach. Educ.* 110:103540. doi: 10.1016/j.tate.2021.103540

Archambault, I., Pascal, S., Tardif-Grenier, K., Dupéré, V., Janosz, M., Parent, S., et al. (2020). The contribution of teacher structure, involvement, and autonomy support on student engagement in low-income elementary schools. *Teach. Teach.* 26, 428–446. doi: 10.1080/13540602.2020.1863208

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., et al. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *Am. Educ. Res. J.* 47, 133–180. doi: 10.3102/0002831209345157

Berger, J.-L., and Girardet, C. (2021). Vocational teachers' classroom management style: the role of motivation to teach and sense of responsibility. *Eur. J. Teach. Educ.* 44, 200–216. doi: 10.1080/02619768.2020.1764930

Breslow, N. E., and Day, N. E. (1987). *Statistical methods in cancer research Volume II. The design and analysis of cohort studies*. Lyon: International Agency for Research on Cancer.

Bruinsma, M., and Jansen, E. P. (2010). Is the motivation to become a teacher related to pre-service teachers' intentions to remain in the profession?. *Eur. J. Teach. Educ.* 33, 185–200. doi: 10.1080/02619760903512927

Centraal Bureau voor de Statistiek [CBS] (2020). *Jaarrapport Integratie*. Available online at: https://longreads.cbs.nl/integratie-2020/ (accessed November 14, 2021).

Chaudron, J., and Dujardin, A. (2019). *Van basisschool tot Universiteit, de werkdruk is totaal uit de klauwen gelopen*. Jacob Bontiusplaats: Trouw.

Den Brok, P., Wubbels, T., and Van Tartwijk, J. (2017). Exploring beginning teachers' attrition in the Netherlands. *Teach. Teach.* 23, 881–895. doi: 10.1080/13540602.2017.1360859

Deng, L., and Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educ. Psychol. Meas.* 77, 185–203. doi: 10.1177/0013164416658325

Dube-Xaba, Z., and Makae, M. P. (2022). Assuring quality of school-based assessment: the perspectives of heads of department in moderating tourism tasks. *Qual. Assur. Educ.* 30, 87–101. doi: 10.1108/QAE-04-2021-0049

Fauth, B., Decristan, J., Decker, A.-T., Büttner, G., Hardy, I., Klieme, E., et al. (2019). The effect of teacher competence on student outcomes in elementary science education: the mediating role of teaching quality. *Teach. Teach. Educ.* 86:102882. doi: 10.1016/j.tate.2019.102882

Gay, G. (2002). Preparing for culturally responsive teaching. *J. Teach. Educ.* 53, 106–116.

Geerlings, J., Thijs, J., and Verkuyten, M. (2018). Teaching in ethnically diverse classrooms: examining individual differences in teacher self-efficacy. *J. Sch. Psychol.* 67, 134–147. doi: 10.1016/j.jsp.2017.12.001

Giersch, J. (2016). A test of personal and social utility values and the appeal of a career in teaching. *Educ. Res. Policy Pract.* 15, 163–173. doi: 10.1007/s10671-016-9194-7

Girardet, C., and Berger, J.-L. (2018). Factors influencing the evolution of vocational teachers' beliefs and practices related to classroom management during teacher education. *Aust. J. Teach. Educ.* 43, 138–158. doi: 10.14221/ajte.2018v43n4.8

Gustafsson, J. E., Nilsen, T., and Hansen, K. Y. (2018). School characteristics moderating the relation between student socioeconomic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Stud. Educ. Eval.* 57, 16–30. doi: 10.1016/j.stueduc.2016.09.004

Harmsen, R., Helms-Lorenz, M., Maulana, R., Van Veen, K., and Van Veldhoven, M. (2019). Measuring general and specific stress causes and stress responses among beginning secondary school teachers in the Netherlands. *Int. J. Res. Method Educ.* 42, 91–108. doi: 10.1080/1743727X.2018.1462313

Heisig, J. P., and Schaeffer, M. (2019). Why you should always include a random slope for the lower-level variable involved in a cross-level interaction. *Eur. Soc. Rev.* 35, 258–279. doi: 10.1093/esr/jcy053

Helms-Lorenz, M., Van de Grift, W., and Maulana, R. (2015). Longitudinal effects of induction on teaching skills and attrition rates of beginning teachers. *Sch. Eff. Sch. Improv.* 27, 178–204. doi: 10.1080/09243453.2015.1035731

Higgins, S., Major, L. E., and Coe, R. (2014). What makes great teaching. *Professional Voice; What Works (and What Doesn't)*. London: Sutton Trust.

Hornstra, L., Mansfield, C., Van der Veen, I., Peetsma, T., and Volman, M. (2015). Motivational teacher strategies: the role of beliefs and contextual factors. *Learn. Environ. Res.* 18, 363–392. doi: 10.1007/s10984-015-9189-y

Huyghebaert, T., Gillet, N., Beltou, N., Tellier, F., and Fouquereau, E. (2018). Effects of workload on teachers' functioning: a moderated mediation model including sleep problems and overcommitment. *Stress Health* 34, 601–611. doi: 10.1002/smi.2820

Inspectie van het Onderwijs (2019). *Technisch rapport leraren en lerarentekort [Technical report teachers and teacher shortage]*. Netherlands: Inspectie van het Onderwijs.

Jennissen, R., Engbersen, G., Bokhorst, M., and Bovens, M. (2018). *De nieuwe verscheidenheid. Toenemende diversiteit naar herkomst in Nederland*. Netherlands: Wetenschappelijke Raad voor het Regeringsbeleid.

Johari, J., Tan, F. Y., Iwani, Z., and Zulkarnain, T. (2018). Autonomy, workload, work-life balance and job performance among teachers. *Int. J. Educ. Manag.* 32, 107–120. doi: 10.1108/IJEM-10-2016-0226

Jones, V. F., and Jones, L. S. (2012). *Comprehensive Classroom Management, Creating Communities of Support and Solving Problems*, 10th Edn. Upper Saddle River: Pearson.

Klassen, R. M., and Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: teacher gender, years of experience, and job stress. *J. Educ. Psychology*102, 741–756. doi: 10.1037/a0019237

Klieme, E. (2013). "The role of large-scale assessments in research on educational effectiveness and school development," in *The Role of International Large-Scale Assessments: Perspectives from Technology, Economy, and Educational Research*, eds M. von Davier, E. Gonzalez, I. Kirsch, and K. Yamamoto (Dordrecht: Springer), 115–147. doi: 10.1007/978-94-007-4629-9

Klieme, E., Pauli, C., and Reusser, K. (2009). "The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms," in *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, eds T. Janik and T. Seidel (Münster: Waxmann), 137–160.

König, J., and Rothland, M. (2012). Motivations for choosing teaching as a career: effects on general pedagogical knowledge during initial teacher education. *Asia Pac. J. Teach. Educ.* 40, 289–315. doi: 10.1080/1359866X.2012.700045

Korpershoek, H., Harms, T., De Boer, H., Van Kuijk, M., and Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Rev. Educ. Res.* 86, 643–680. doi: 10.3102/0034654315626799

Krammer, G., Pflanzl, B., and Mayr, J. (2019). Using students' feedback for teacher education: measurement invariance across pre-service teacher-rated and student-rated aspects of quality of teaching. *Assess. Eval. High. Educ.* 44, 596–609. doi: 10.1080/02602938.2018.1525338

Krieg, J. M. (2006). Teacher quality and attrition. *Econ. Educ. Rev.* 25, 13–27. doi: 10.1016/j.econedurev.2004.09.004

Lazarides, R., Fauth, B., Gaspard, H., and Göllner, R. (2021). Teacher self-efficacy and enthusiasm: relations to changes in perceived teaching quality at the beginning of secondary education. *Learn. Instruct.* 73:101435. doi: 10.1016/j.learninstruc.2020.101435

Liu, S., Keeley, J. W., and Sui, Y. (2020). Multilevel of factors influencing teacher job satisfaction in China: evidence from the TALIS 2018. *Educ. Stud.* doi: 10.1080/03055698.2020.1837615

Maulana, R., Opdenakker, M. C., and Bosker, R. (2016). Teachers' instructional behaviors as important predictors of academic motivation: changes and links across the school year. *Learn. Individ. Differ.* 50, 147–156. doi: 10.1016/j.lindif.2016.07.019

McGrath-Champ, S., Wilson, R., Stacey, M., and Fitzgerald, S. (2018). *Understanding work in schools: The foundation for teaching and learning. 2018 Report to the NSW Teachers Federation*. Sydney: NSW Teachers Federation.

Nuffic De Nederlandse organisatie voor internationalisering in onderwijs. Available online at: https://www.nuffic.nl/en/education-systems/netherlands/primary-and-secondary-education (accessed February 9, 2022).

O'Brien, L. M., and Schillaci, M. (2002). Why do I want to teach, anyway? Utilizing autobiography in teacher education. *Teach. Educ.* 13, 25–40. doi: 10.1080/1047210120128564

Organisation for Economic Co-Operation and Development [OECD] (2019). *TALIS 2018 Technical Report*. Paris: OECD Publishing.

Parkhouse, H., Lu, C. Y., and Massaro, V. R. (2019). Multicultural education professional development: a review of the literature. *Rev. Educ. Res.* 89, 416–458.

Parr, A., Gladstone, J., Rosenzweig, E., and Wang, M.-T. (2021). Why do I teach? A mixed-methods study of in-service teachers' motivations, autonomy-supportive instructions, and emotions. *Teach. Teach. Educ.* 98:103228. doi: 10.1016/j.tate.2020.103228

Patall, E. A., and Zambrano, J. (2019). Facilitating student outcomes by supporting autonomy: implications for practice and policy. *Policy Insights Behav. Brain Sci.* 6, 115–122. doi: 10.1177/2372732219862572

Payne, C. M. (2014). *So much reform, so little change. The persistence of failure in urban schools*. Cambridge: Harvard University Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88, 879–903. doi: 10.1037/0021-9010.88.5.879

Richardson, P. W., and Watt, H. G. (2014). "Why people choose teaching as a career: An expectancy-value approach to understanding teacher motivation," in *Teacher Motivation: Theory and Practice*, eds P. W. Richardson, S. A. Karabenick, and H. M. G. Watt (New York: Routledge), 3–19. doi: 10.4324/9780203119273-1

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., and Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: the mediating role of instructional quality. *Learn. Instr.* 32, 63–72. doi: 10.1016/j.learninstruc.2014.01.007

Sapulete, S., De Wilde, P., Ouwehand, K., and Van der Boom, E. (2018). *Teaching and Learning International Survey (TALIS) 2018. Nationaal rapport voortgezet onderwijs. [National report secondary education]*. Netherlands: Rijksoverheid.

Scheerens, J. (1990). School effectiveness research and the development of process indicators for school functioning. *Sch. Eff. Sch. Improv. J.* 1, 61–80. doi: 10.1080/0924345900010106

Scheerens, J. (2000). *Improving school effectiveness. Fundamentals of Educational Planning (No. 68)*. Paris: United Nations Educational, Scientific and Cultural Organization.

Scherzinger, M., and Wettstein, A. (2019). Classroom disruptions, the teacher–student relationship and classroom management from the perspective of teachers, students and external observers: a multimethod approach. *Learn. Environ. Res.* 22, 101–116. doi: 10.1007/s10984-018-9269-x

Severiens, S., Wolff, R., and Van Herpen, S. (2014). Teaching for diversity: a literature overview and an analysis of the curriculum of a teacher training college. *Eur. J. Teach. Educ.* 37, 295–311. doi: 10.1080/02619768.2013.845166

Thijs, J., Westhof, S., and Koomen, H. (2012). Ethnics incongruence and the student-teacher relationship: the perspective of ethnic majority teachers. *J. Sch. Psychol.* 50, 257–273. doi: 10.1016/j.jsp.2011.09.004

Tielman, K., Wesselink, R., and Den Brok, P. (2021). Tensions experienced by teachers of Dutch culturally diverse senior secondary vocational education and training: an exploratory study. *Int. J. Train. Dev.* 1–18. doi: 10.1111/ijtd.12238

Traag, T., (2018). *Statistische Trends: Leerkrachten in het Basisonderwijs*. Netherlands: Centraal Bureau voor de Statistiek.

Vereninging van Scholen in het Voortgezet Onderwijs [VO RAAD] (n.d.). *Lerarentekort*. Available online at: https://www.vo-raad.nl/themas/ontwikkeling-leraren/onderwerpen/lerarentekort-e38d9309-3dfb-4caf-990a-1c0ad2d0136d (accessed November 15, 2021)

Watt, H. M., and Richardson, P. W. (2007). Motivational factors influencing teaching as a career choice: development and validation of the FIT-Choice scale. *J. Exp. Educ.* 75, 167–202. doi: 10.1016/j.tate.2012.03.003

Watt, H. M., and Richardson, P. W. (2012). An introduction to teaching motivations in different countries: comparisons using the FIT-Choice scale. *Asia Pac. J. Teach. Educ.* 40, 185–197. doi: 10.1080/1359866X.2012.700049

Wentzel, K. R. (2009). "Students' relationships with teachers as motivation contexts," in *Handbook of Motivation in School*, eds K. Wentzel and A. Wigfield (Mahwah: Erlbaum), 301–322.

Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teach. Coll. Rec.* 112, 1008–1037. doi: 10.1177/016146811011200408

Wubbels, T., Den Brok, P., Veldman, I., and Van Tartwijk, J. (2006). Teacher interpersonal competence for Dutch secondary multicultural classrooms. *Teach. Teach. Theor. Pract.* 12, 407–433. doi: 10.1080/1315060060061126

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

APPENDIX

Items of All Scales Used in the Analyses

Social Utility Motivation to Teach (T3SOCUT)

TT3G07: How important were the following for you to become a teacher?
Response options: "Not important at all" (1), "Of low importance" (2), "Of moderate importance" (3), "Of high importance" (4).
TT3G07E Teaching allowed me to influence the development of children and young people.
TT3G07F Teaching allowed me to benefit the socially disadvantaged.
TT3G07G Teaching allowed me to provide a contribution to society.

Personal Utility Motivation to Teach (T3PERUT)

TT3G07: How important were the following for you to become a teacher?
Response options: "Not important at all" (1), "Of low importance" (2), "of moderate importance" (3), "Of high importance" (4).
TT3G07A Teaching offered a steady career path.
TT3G07B Teaching provided a reliable income.
TT3G07C Teaching was a secure job.
TT3G07D The teaching schedule (e.g., hours, holidays, part-time positions) fit with responsibilities in my personal life.

Workload Stress (T3WLOAD)

TT3G52: Thinking about your job at this school, to what extent are the following sources of stress in your work?
Response options: "Not at all" (1), "To some extent" (2), "Quite a bit" (3), "A lot" (4).
TT3G52A Having too much lesson preparation.
TT3G52B Having too many lessons to teach.
TT3G52C Having too much marking.
TT3G52D Having too much administrative work to do (e.g., filling out forms).
TT3G52E Having extra duties due to absent teachers.

Clarity of Instruction (T3CLAIN)

TT3G42: Thinking about your teaching in the <target class>, how often do you do the following?
Response options: "Never or almost never" (1), "Occasionally" (2), "Frequently" (3), "Always" (4).
TT3G42A I present a summary of recently learned content.
TT3G42B I set goals at the beginning of instruction.
TT3G42C I explain what I expect the students to learn.
TT3G42D I explain how new and old topics are related.

Classroom Management (T3CLASM)

TT3G42: Thinking about your teaching in the <target class>, how often do you do the following?
Response options: "Never or almost never" (1), "Occasionally" (2), "Frequently" (3), "Always" (4).
TT3G42I I tell students to follow classroom rules.
TT3G42J I tell students to listen to what I say.
TT3G42K I calm students who are disruptive.
TT3G42L When the lesson begins, I tell students to quieten down quickly.

Cognitive Activation (T3COGAC)

TT3G42: Thinking about your teaching in the <target class>, how often do you do the following?
Response options: "Never or almost never" (1), "Occasionally" (2), "Frequently" (3), "Always" (4).
TT3G42E I present tasks for which there is no obvious solution.
TT3G42F I give tasks that require students to think critically.
TT3G42G I have students work in small groups to come up with a joint solution to a problem or task.
TT3G42H I ask students to decide on their own procedures for solving complex tasks.

Student-Teacher Relations (T3STUD)

TT3G49: How strongly do you agree or disagree with the following statements about what happens in this school?
Response options: "Strongly disagree" (1), "Disagree" (2), "Agree" (3), "Strongly agree" (4).
TT3G49A Teachers and students usually get on well with each other.
TT3G49B Most teachers believe that the students' well-being is important.
TT3G49C Most teachers are interested in what students have to say.
TT3G49D If a student needs extra assistance, the school provides it.

frontiers | Frontiers in Psychology

ORIGINAL RESEARCH
published: 23 May 2022
doi: 10.3389/fpsyg.2022.853221

Chinese Students' Perceptions of the Motivational Climate in College English Courses: Relationships Between Course Perceptions, Engagement, and Achievement

Ming Li[1], Brett D. Jones[2], Thomas O. Williams[2] and Yingjian Guo[3]*

[1]School of Foreign Languages, Shanghai University of Engineering Sciences, Shanghai, China, [2]School of Education, Virginia Tech, Blacksburg, VA, United States, [3]School of Foreign Languages, Renmin University of China, Beijing, China

OPEN ACCESS

Edited by:
Ridwan Maulana, University of Groningen, Netherlands

Reviewed by:
Blanka Klimova, University of Hradec Králové, Czechia
Tomislav Mestrovic, University North, University Centre Varazdin, Croatia

***Correspondence:**
Brett D. Jones
brettjones@vt.edu

Specialty section:
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

Received: *12 January 2022*
Accepted: *19 April 2022*
Published: *23 May 2022*

Citation:
Li M, Jones BD, Williams TO and Guo Y (2022) Chinese Students' Perceptions of the Motivational Climate in College English Courses: Relationships Between Course Perceptions, Engagement, and Achievement. Front. Psychol. 13:853221. doi: 10.3389/fpsyg.2022.853221

Effective teachers create a motivational climate that engages students in course activities in ways that lead to increased learning and achievement. Although researchers have identified motivational climate variables that are associated with students' engagement and achievement, less is known about how these variables are related in different courses and cultures. The purpose of the two studies presented in this paper was to contribute to this research literature by examining these associations within the context of college English courses in two Chinese universities. Specifically, we investigated the relationships between students' perceptions of the motivational climate (i.e., perceptions of empowerment/autonomy, usefulness, success, interest, and caring), cognitive and behavioral engagement, and achievement. This is the first study to examine the connections between all of these variables in one path model in college English courses in China. We administered surveys at two different Chinese universities ($n = 332$ and 259) and used regression and path analysis to examine the relationships among the variables. We demonstrated that (a) students' perceptions of the motivational climate were related to their cognitive engagement, (b) cognitive engagement was related to their behavioral engagement, and (c) behavioral engagement predicted their achievement. These findings are consistent with and extend the growing body of literature on motivational climate and engagement, and they highlight the importance of some motivational climate perceptions over others as significant predictors of cognitive engagement. We conclude that effective English language teachers in China do the following: help students to believe that they can be successful, trigger and maintain students' interest, and empower students by providing them with choices in activities and assignments.

Keywords: motivation, engagement, MUSIC Model of Motivation, English courses, course perceptions, motivational climate, foreign language instruction, English as a second language

INTRODUCTION

Effective teachers engage students in course activities (Christenson et al., 2012b), which is important because students' engagement is related to their achievement (Muenks et al., 2017; Tao et al., 2022). Researchers have identified a variety of factors associated with student engagement in courses, including their perceptions of the motivational climate (see Christenson et al., 2012b; Jones et al., 2021). Yet, the role of engagement as a link between students' motivational climate perceptions and their achievement remains unclear, in part, because researchers do not always agree on the exact definitions of engagement or the order in which different types of engagement occur (Reschly and Christenson, 2012). For example, it has been suggested that cognitive engagement precedes behavioral engagement (Reschly and Christenson, 2012; Reeve et al., 2020), and this suggestion has been documented empirically by researchers (Jones and Carter, 2019).

The purpose of the present studies was to examine associations between university students' motivational climate perceptions, engagement, and achievement in English language (EL) courses in China. We chose this context for a couple reasons. All Chinese college students are required to enroll in college EL courses and complete national English tests (Guo et al., 2020). Thus, understanding how students' perceptions of the motivational climate in EL courses are related to their engagement in these courses—and subsequently achievement in these courses—could be useful to the large number of instructors who teach these courses. In addition, despite the importance of college EL courses, many Chinese students are unmotivated in these courses (Li et al., 2016; Li, 2021) and lack the skills needed to pass the national exams (Hertling, 1996). Understanding how students' motivation-related perceptions in EL courses affect their engagement and achievement could lead to the development of effective instructional strategies and interventions aimed at engaging students.

More specifically, the present studies can contribute to the literature about effective teaching in two ways. First, the results will determine whether the motivational climate constructs that have been shown to affect students' engagement in other contexts also affect students' engagement in EL courses in China. Relatedly, the results will identify which motivational climate constructs are most salient in this context. Second, the results will determine whether the motivational climate constructs and engagement can be linked to achievement in this context; and if so, whether cognitive engagement precedes behavioral engagement as hypothesized. Together, these findings will provide a clearer understanding of the motivational climate factors that can affect students' engagement in EL courses in China. Teachers can then focus on incorporating teaching strategies related to these factors in order to increase students' engagement and achievement.

LITERATURE REVIEW

Motivation and Engagement in Courses

Engagement is a broad psychological construct that has multiple definitions and has been studied in a variety of contexts. Many researchers consider motivation to precede engagement and define motivation as one's intentions to engage (Jones, 2018, 2020b) and engagement as one's actions (Christenson et al., 2012a). Engagement can be further divided into a few dimensions, including behavioral engagement (e.g., effort, actual participation in school and learning), cognitive engagement (e.g., cognitive investment in the coursework, such as mental effort and use of effective learning strategies), and emotional engagement (e.g., students' emotional responses to teachers, peers, and the school environment, such as enjoyment and anxiety; Fredricks et al., 2004). Students' engagement predicts many different positive educational outcomes, such as achievement, learning, and the likelihood of high school completion (Reschly and Christenson, 2012; Skinner and Pitzer, 2012; Tao et al., 2022).

Many different psychological theories have been used to explain students' engagement in educational settings, such as self-determination theory (Reeve, 2012; Ryan and Deci, 2020), social cognitive theory (Bandura, 1997; Schunk and DiBenedetto, 2020), self-regulation theories (Cleary and Zimmerman, 2012), theories of emotions (Pekrun and Linnenbrink-Garcia, 2012), and interest theories (Ainley, 2012; Renninger and Hidi, 2015; for more perspectives, see Christenson et al., 2012b). Simultaneously, a mostly separate research literature has developed to explain the motivation of students in second/foreign language (L2) courses (Al-Hoorie, 2017; Al-Hoorie and MacIntyre, 2019). This research has led to notable contributions such as the Socio-educational Model of Second Language Acquisition by Gardner (2019) and the L2 motivational self-system by Dörnyei (2009). Although L2 researchers have made some connections between teachers' motivational strategies, students' motivation, students' engagement, and students' achievement (e.g., Alrabai, 2016), the research directly related to students' perceptions in a course and the effects of these perceptions on students' engagement and achievement has been limited (see Lamb, 2019, for a review). Instead, L2 researchers have focused on studying the motivation of students, as opposed to focusing on motivating students in courses (Boo et al., 2015; Lamb, 2019).

Recently, Jones (2020a) has suggested that the MUSIC Model of Motivation (abbreviated in this paper as the MUSIC model; Jones, 2009, 2018) could be applied to L2 instruction in a manner that "does not replace existing L2 motivation theories, but rather…used as a complementary approach" (p. 2). The multidimensional MUSIC model highlights five student perceptions of the motivational climate—perceptions that can be linked to current motivation-related constructs and psychological theories—that have been shown to be associated with student engagement (Jones, 2010, 2019; Jones et al., 2021) and course ratings (Wilkins et al., 2021; Jones et al., 2022b). The five motivational climate perceptions include students' perceptions of: their autonomy/empowerment in the class (*eMpowerment*), the usefulness/utility value of the content and activities (*Usefulness*), the extent to which they can be successful if they put forth effort (*Success*), their enjoyment/interest during the activity (*Interest*), and whether the instructor and other students care about their learning and about them personally (*Caring*; the five keywords form the acronym MUSIC: eMpowerment, Usefulness, Success, Interest, and Caring). Evidence that these perceptions affect students' motivation and

engagement is provided by a variety of theories, including arousal theories (Berlyne, 1960), attachment theory (Bowlby, 1969), attribution theory (Weiner, 2000), situated expectancy-value theory (Eccles and Wigfield, 2020), interest theories (Schraw and Lehman, 2001; Renninger and Hidi, 2015), self-determination theory (Ryan and Deci, 2020), social cognitive theory (Bandura, 1997; Schunk and DiBenedetto, 2020), among others (see Jones, 2018 for a more comprehensive list). Examples of empirical research studies linking these five motivational climate perceptions and student engagement include the following: Giving students' autonomy (empowerment) has been shown to increase student engagement (Reeve et al., 2004; Jang et al., 2012), perceived instrumentality (usefulness) and self-efficacy (success) were found to increase engagement (Walker and Greene, 2009), curiosity (situational interest) has been associated with students engagement in science (Wu and Wu, 2020), and several studies have shown that the caring relationship between a teacher and students leads to higher student engagement and achievement (King, 2015; Quin, 2016). Thus, the MUSIC model constructs are consistent with current psychological constructs and theories. In addition, the MUSIC constructs are consistent with motivational teaching strategies proposed by L2 researchers (e.g., Dörnyei and Ushioda, 2011), such as the need for instructors to: support learner autonomy (empowerment), help students to connect the relevance of course activities to their lives (usefulness), increase learners' expectancy of success (success), get students interested in course activities, and foster relationships between teachers and students (caring).

MUSIC Perceptions and Engagement

Figure 1 shows how, in the MUSIC model, external and internal variables affect students' perceptions of the motivational climate in a course, which then affects their motivation, engagement, and learning/performance. Although students have a variety of perceptions within a course, their MUSIC perceptions (i.e., perceptions of empowerment, usefulness, success, interest, and caring) have been studied because they relate to important outcomes and have been shown to be distinct; that is, they refer to different constructs that are separable through factor analyses (Mohamed et al., 2013; Jones et al., 2014, 2016, 2019; Jones and Sigmon, 2016; Pace et al., 2016; Schram and Jones, 2016; Chittum and Jones, 2017; Tendhar et al., 2017; Gladman et al., 2020).

The MUSIC perceptions have been shown to be correlated to their students' engagement in studies of undergraduate students (Jones, 2010, 2019; Jones et al., 2021), and several decades of research have documented that constructs related to these perceptions are related to students' engagement (for reviews, see Christenson et al., 2012a; Schunk et al., 2014). Yet, studies that have included all five MUSIC perceptions sometimes find that some, but not all, of the MUSIC perceptions are related to student engagement in different educational contexts. For example, in a study of college courses, Jones (2019) documented that all five MUSIC perceptions were related to students' behavioral engagement in some courses, but that only two, three, or four of the MUSIC perceptions were related to behavioral engagement in other courses. Within EL courses in China, only one study (Li et al., 2016) has examined the relationship between all five MUSIC perceptions and students' engagement. This study showed that although empowerment, usefulness, success, and interest (but not caring) were correlated with student engagement, only empowerment and success were significantly related to engagement when all of these variables were included in one regression model to predict engagement. This study was limited to 101 students in three classes at one college; and therefore, more studies replicating this study in other contexts would be useful.

In addition, only two studies (Jones and Carter, 2019; Jones et al., 2021) have included constructs related to all five MUSIC perceptions, engagement, and learning/achievement. The Jones and Carter study was conducted within a psychology course at a university in the United States and demonstrated that while all five MUSIC constructs were significantly correlated with students' cognitive and behavioral engagement, only empowerment and usefulness were significantly related to cognitive engagement when all the constructs were included in one model. Furthermore, cognitive and behavioral engagement were significantly related to student learning; and importantly, the MUSIC constructs predicted cognitive engagement, which was modeled to precede behavioral engagement and then learning. In Jones et al. (2021) study, the MUSIC perceptions of students in an online geography course at a United States university were positively correlated with their effort in the course. When all five MUSIC constructs were included in a structural equation model with MUSIC perceptions predicting behavioral engagement, only interest and caring were significant predictors of behavioral engagement, which then predicted achievement. Findings from studies such as these demonstrate that different course perceptions can be more influential in some courses than others. Relationships between these variables need to be better understood because if patterns are found in EL courses in China (e.g., usefulness is most strongly related to engagement and achievement), they could be used to help instructors to design courses that target these perceptions to lead to increased student engagement and achievement. Therefore, the purpose of the present studies was to identify the MUSIC perceptions that are most highly associated with students' engagement in EL courses in China.

PURPOSE AND RESEARCH QUESTIONS

More research is needed within specific contexts (e.g., different types of courses at different colleges) to document which MUSIC perceptions are most salient in different contexts. For example, empowerment and usefulness may be most important to engaging students in a psychology course in the United States (as demonstrated in the study by Jones and Carter, 2019), but perceptions of empowerment and success might be more important in a college English language course in China (as demonstrated in the study by Li et al., 2016). The aim of the present studies was to investigate the effects of students' MUSIC perceptions in the context of college EL courses in China to

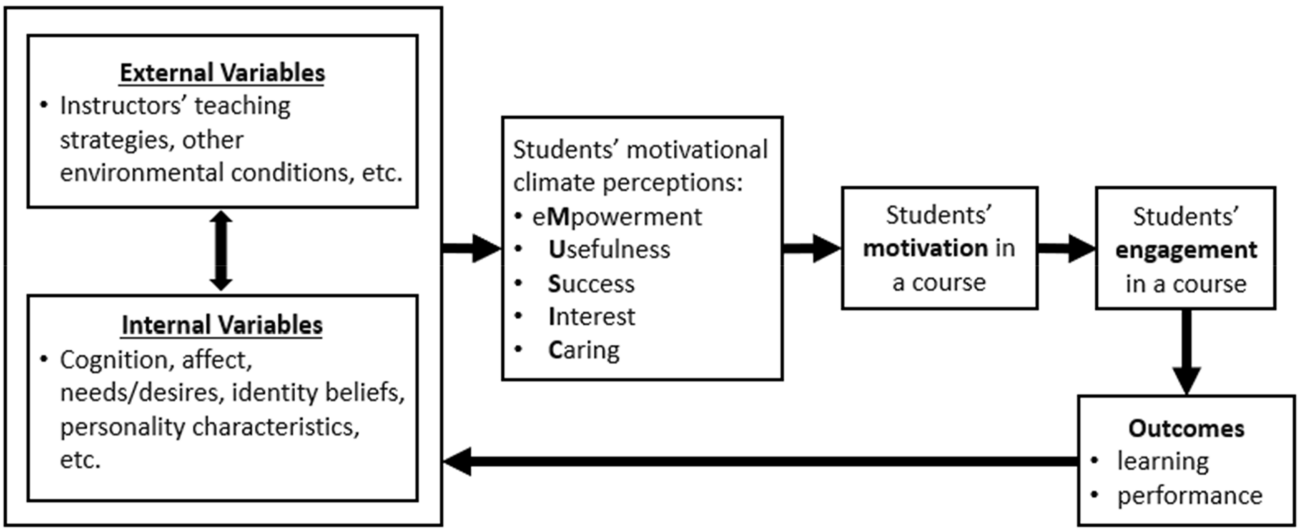


FIGURE 1 | Key elements of the MUSIC model of motivation. Adapted from "Motivating Students by Design: Practical Strategies for Professors" by Jones (2018). Used with permission.

better understand (a) which MUSIC perceptions are most important in predicting cognitive and behavioral engagement in EL courses in China, (b) the extent to which cognitive and behavioral engagement in EL courses are related to EL achievement, and (c) whether cognitive engagement precedes behavioral engagement as mediators between MUSIC perceptions and achievement. The results of these studies could help EL instructors to become more effective by identifying the components of the motivational climate that are most important in engaging students in their courses. Teachers could then design instructional activities to support these components of the motivational climate.

In the present studies, we chose to focus on cognitive and behavioral engagement instead of emotional engagement for a few reasons. First, emotional engagement overlaps conceptually with situational interest, which is one of the course perceptions already included in the MUSIC model. In the MUSIC model, situational interest is defined similar to how other researchers have defined it, as "liking and willful engagement in a cognitive activity" (Schraw and Lehman, 2001, p. 23). This definition includes an affective component (the liking) similar to emotional engagement—which includes students' emotional responses such as enjoyment (Fredricks et al., 2004)—and includes a willful engagement. Reeve et al. (2020) provided empirical evidence of the conceptual overlap between emotional engagement and interest/intrinsic motivation. Second, emotional engagement likely precedes engagement or "amplifies on-going and future behavioral engagement" (Reeve et al., 2020, p. 8). And finally, emotional engagement does not consistently predict educational outcomes such as achievement (Gutiérrez and Tomás, 2019; Reeve et al., 2020).

We conducted two studies to examine the extent to which students' MUSIC perceptions in an English language course in China affect their cognitive engagement, behavioral engagement, and achievement. Our five specific research questions were as follows:

- RQ1 (Study 1 and Study 2): To what extent do students' MUSIC perceptions relate to their cognitive engagement?
- RQ2 (Study 1 and Study 2): To what extent do students' MUSIC perceptions relate to their behavioral engagement?
- RQ3 (Study 2): To what extent do students' MUSIC perceptions positively predict their cognitive engagement, which then positively predicts their achievement?
- RQ4 (Study 2): To what extent do students' MUSIC perceptions positively predict their behavioral engagement, which then positively predicts their achievement?
- RQ5 (Study 2): To what extent do students' MUSIC perceptions positively predict their cognitive engagement, which then positively predicts their behavioral engagement, which then positively predicts their achievement?

We conducted Study 1 to examine RQ1 and RQ2 as a proof of concept that one or more MUSIC perceptions were related to students' cognitive and behavioral engagement in college English courses in China. After providing evidence of these relationships, we conducted Study 2 with a different sample of students to provide evidence that the results of Study 1 were generalizable to students attending another university and to answer RQ3, RQ4, and RQ5 by modeling the relationships between MUSIC perceptions, cognitive and behavioral engagement, and achievement.

We predicted that students' MUSIC perceptions would be positively related to their cognitive and behavioral engagement based on studies that have documented these relationships with non-EL courses in the United States (e.g., Jones, 2019; Jones and Carter, 2019; Jones et al., 2021), students with non-EL courses in China (Jones et al., 2017), and studies with EL courses in China (Li et al., 2016). We also predicted that

students' cognitive and behavioral engagement would positively relate to their achievement based on studies that have shown these relationships with non-EL courses in the United States (e.g., Muenks et al., 2017; Di Leo et al., 2019; Jones and Carter, 2019). Our third prediction was that cognitive and behavioral engagement would mediate the relationship between students' MUSIC perceptions and their achievement. We based this prediction on the reasoned hypotheses by some researchers (Reschly and Christenson, 2012; Reeve et al., 2020) and the empirical findings of others (Jones and Carter, 2019).

STUDY 1

Method

Participants

Participants in Study 1 were 332 undergraduate students at a large university in northeastern China. The students were enrolled in one of four English classes ($n = 46$, 76, 100, 110) for non-English majors. More of the participants were female ($n = 250$; 75.3%) than male ($n = 82$; 24.7%). The majority of the participants reported that they were Han nationality ($n = 203$, 61.1%), whereas 127 participants (38.3%) reported that they were one of the minority nationalities, and two students (0.6%) reported that they were an "other" race/ethnicity. Regarding their class standing, 77 (23.2%) were first-year students, 254 (76.5%) were sophomores, and 1 (0.3%) was a senior. Most students ($n = 327$; 98.5%) were between the ages of 18 and 22, and five students (1.5%) were older than 22 years old. Most or all students had been enrolled in English classes in school for at least 10 years (since elementary school).

Procedure

Students completed an online questionnaire near the end of the semester in their English course. Because all the participants were learning English (and thus, not fluent in English), all of the questionnaire items were provided in Chinese. Students completed the survey as part of their normal class activities; and therefore, consent to participate in this study was not obtained. Instead, we received the anonymous data from the instructors as existing data and their inclusion in this study was approved as "Exempt" by the authors' Institutional Review Board (IRB #17-021).

Motivational Climate

We measured the motivational climate using the MUSIC® Model of Academic Motivation Inventory (College Student short-form version; available at Jones, 2012/2021), which consists of 20 items that form five scales: a four-item empowerment scale (measuring autonomy; Ryan and Deci, 2020), a four-item usefulness scale (measuring utility value; Eccles and Wigfield, 2020), a four-item success scale (measuring expectancy for success; Eccles and Wigfield, 2020), a four item interest scale (measuring situational interest; Renninger and Hidi, 2015), and a four-item caring scale (measuring caring; Noddings, 1992). All items were responded to on a six-point Likert-format scale: 1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Somewhat disagree*, 4 = *Somewhat agree*, 5 = *Agree*, 6 = *Strongly agree*. Example items include: "I had flexibility in what I was allowed to do in this course" (empowerment), "In general, the coursework was useful to me" (usefulness), "I was confident that I could succeed in the coursework" (success), "The coursework was interesting to me" (interest), and "The instructor cared about how well I did in this course" (caring). The MUSIC Inventory produces reliable and valid scores and correlates with other measures as expected (Jones and Skaggs, 2016; Pace et al., 2016; Jones et al., 2019). The Chinese translation of the MUSIC Inventory has also been shown to demonstrate acceptable psychometric properties; for example, Cronbach's alpha values for the scales were 0.82 for empowerment, 0.89 for usefulness, 0.87 for success, 0.93 for interest, and 0.88 for caring (Jones et al., 2017).

Behavioral Engagement

To measure behavioral engagement, we used a three-item effort scale that was based on the Effort/Importance scale, which is part of the Intrinsic Motivation Inventory (Ryan, 1982; McAuley et al., 1989). All items included a 6-point Likert-format scale (1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Somewhat disagree*, 4 = *Somewhat agree*, 5 = *Agree*, 6 = *Strongly agree*). An example item is: "I put a lot of effort into this course." In Jones (2010), the reliability estimates were good ($\alpha = 0.84$, 0.84, 0.86, 0.84). We used the Chinese translation of this scale that was used in Jones et al. (2017).

Cognitive Engagement

We used the 8-item Self-Regulated Strategy Use scale that is part of the Student Perceptions of Classroom Knowledge-Building Scale (SPOCK; Shell et al., 2005; Shell and Husman, 2008) to assess cognitive engagement. The Self-Regulated Strategy Use scale measures the extent to which students' behaviors and strategies are associated with self-regulation (e.g., planning, goal setting, monitoring, and evaluation of studying and learning). An example item is: "I try to determine the best approach for studying each assignment" (1 = *Almost never*, 2 = *Seldom*, 3 = *Sometimes*, 4 = *Often*, 5 = *Almost always,* 6 = *Always*). Reliability estimates have been shown to be acceptable ($\alpha = 0.81$; Shell and Husman, 2008). We used the Chinese translation of this scale that was used in Jones et al. (2017).

Analysis

We used IBM® SPSS® version 26 to compute measures of dispersion, distribution, and correlation, and to conduct an exploratory factor analysis (EFA) to examine the psychometric properties of the MUSIC Inventory items. We used Amos version 25 to run regression analyses with the MUSIC constructs predicting cognitive engagement in one model and behavioral engagement in another. For all statistical tests, we set the alpha value at 0.05, and we report two-tailed values of p.

Results

We conducted an EFA on the 20-item MUSIC Inventory using a principal factors analysis and a Promax rotation with Kaiser normalization ($n = 332$). We removed two items (an empowerment

item and a usefulness item) because they loaded higher on a factor other than their intended factor. All the other items loaded on their factors as anticipated. Only three items cross-loaded on another factor at value greater than 0.20. We retained these three items because (a) the cross-loading values (i.e., 0.24, 0.32, and 0.34) were not very high (Tabachnick and Fidell, 2007), (b) the items have been shown to load on these factors in other studies (e.g., Jones et al., 2019), and (c) the items have high face validity (e.g., the item "The coursework is interesting to me" asks students about "interest"). The five factors explained 73.3% of the variance. The 0.0000106 value for the determinant of the correlation matrix was acceptable (Field, 2000), the 0.93 value for the Kaiser-Meyer-Olkin measure of sampling adequacy was "marvelous" (Kaiser, 1974), and the Bartlett test of sphericity was statistically significant ($\chi^2[153] = 3{,}711.9$, $p < 0.001$).

The Cronbach's alpha values for all the measured variables were acceptable (George and Mallery, 2003), ranging from 0.79 to 0.93 (see **Table 1**). Descriptive statistics for each variable and correlations between the study variables are also provided in **Table 1**. The correlations among the MUSIC variables ranged from 0.48 to 0.73 and correlations between the MUSIC variables and behavioral and cognitive engagement, varied from 0.33 to 0.61.

We conducted regressions with students' MUSIC perceptions predicting cognitive engagement in one model (**Figure 2**) and behavioral engagement in the other model (**Figure 3**). Success and interest were significant predictors of both cognitive and behavioral engagement. Empowerment was also a significant predictor of cognitive engagement. The MUSIC constructs explained 39.0% of the variance in cognitive engagement and 36.4% of the variance in behavioral engagement.

Discussion

Our first two research questions led us to examine the extent to which students' MUSIC perceptions were related to their cognitive engagement (RQ1) and behavioral engagement (RQ2). The correlations in **Table 1** and significant paths in **Figure 2** (Model 1a) and **Figure 3** (Model 1b) indicate that students' MUSIC perceptions were significantly related to both their cognitive and behavioral engagement. The reason we included all five of the MUSIC variables in one regression model for Models 1a and 1b was to determine whether some of the motivational climate variables were more important than others in predicting cognitive and behavioral engagement. Success, interest, and empowerment were statistically significant predictors of cognitive engagement, whereas success and interest were statistically significant predictors of behavioral engagement.

These findings provided evidence that students' MUSIC perceptions are related to their cognitive and behavioral engagement in college English courses in China. Given these findings, we conducted a second study to determine whether these results were generalizable to students at a different university. Furthermore, we wanted to examine whether students' cognitive and behavioral engagement predicted their achievement (RQ3, RQ4, and RQ5).

STUDY 2

Method

Participants

Participants in Study 2 were 259 undergraduates at a large university in mid-eastern China. The students were enrolled in one of six English classes (n was about 45 students per class) for non-English majors. More of the participants were female ($n = 180$; 69.5%) than male ($n = 79$; 30.5%). Most of the participants reported that they were Han nationality ($n = 255$, 98.5%), whereas four participants (1.5%) reported they were one of the minority nationalities. All the participants were 1st-year students who ranged in age from 18 to 21. Most or all students had been enrolled in English classes in school for at least 10 years (since elementary school).

Procedure

The procedures were similar to those in Study 1 except that students completed the questionnaire with paper-and-pencil (instead of online) near the end of the semester of their English course. Also, students gave their written consent to participate in the study prior to completing the questionnaire. These data were approved for inclusion in this study by the authors' Institutional Review Board (IRB #16-932).

TABLE 1 | Correlations among Study 1 variables.

	1	2	3	4	5	6	7
1. Empowerment							
2. Usefulness	0.48						
3. Success	0.69	0.55					
4. Interest	0.68	0.60	0.73				
5. Caring	0.56	0.57	0.52	0.58			
6. Behavioral engagement	0.47	0.37	0.55	0.56	0.33		
7. Cognitive engagement	0.51	0.37	0.61	0.52	0.33	0.68	
M	4.71	5.00	4.38	4.47	5.31	4.19	3.70
SD	0.85	0.80	0.93	0.96	0.60	1.09	1.02
Cronbach's α	0.79	0.79	0.87	0.89	0.82	0.92	0.93

$p < 0.001$ for all correlations.

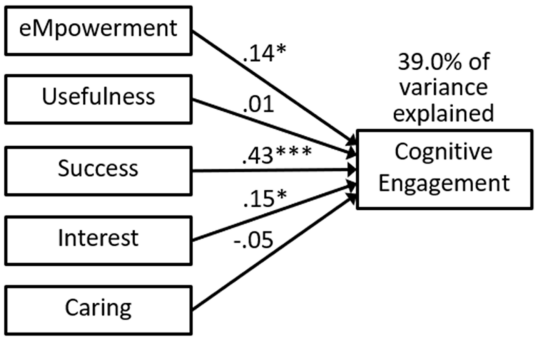


FIGURE 2 | Model 1a regression predicting cognitive engagement. Statistics are standardized beta coefficients. * $p < 0.05$, *** $p < 0.001$.

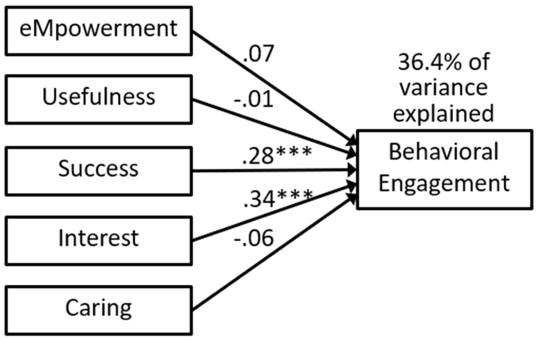


FIGURE 3 | Model 1b regression predicting behavioral engagement. Statistics are standardized beta coefficients. *** $p < 0.001$.

Measures

All the measures were the same as those used in Study 1. In addition, we used a measure of achievement that included students' scores from the final English test that was administered at the end of their first college English course. This test was developed by instructors at another university to resemble (in structure and content) the national English test that students take after completing three courses in college English. Therefore, some of the content on the exam had not been covered in the course in which students were currently enrolled because this course was only their first college English course. The test included a writing section (students wrote an essay), a listening section (students listened to daily news and conversation in English and answered multiple-choice questions), a reading comprehension section (students read passages and then inserted paragraphs into the correct position in an article and answered multiple-choice and fill-in-the-blank questions), and a translation section (students translated a paragraph from Chinese to English). The range of possible test scores was 0–100.

Analysis

We used IBM® SPSS® version 26 to compute measures of dispersion, distribution, and correlation. IBM® SPSS® Amos™ version 25 was used to estimate the structural models with students' MUSIC perceptions predicting cognitive engagement, behavioral engagement, and/or achievement as described in the "Results" section. We assessed the construct validity of the 18-item MUSIC Inventory used in Study 1 by conducting a confirmatory factor analysis (CFA). For all statistical tests, we set the alpha value at 0.05, and we report two-tailed values of p.

Results

The results of the CFA for the 18-item MUSIC Inventory are presented here: $\chi^2 = 289.82$, $df = 125$, $p < 0.001$; and the fit indices indicated a reasonable fit to the data (Hu and Bentler, 1999; Kline, 2005), with CFI = 0.911, SRMR = 0.063, and RMSEA = 0.071 (90% confidence interval ranged from 0.061 to 0.082). In addition, the Cronbach's alpha values for the MUSIC constructs, cognitive engagement, and behavioral engagement ranged from acceptable to excellent (see **Table 2**; George and Mallery, 2003). Descriptive statistics for each variable and correlations between the study variables are also provided in **Table 2**. The correlations among the MUSIC variables ranged from 0.22 to 0.59, which are slightly lower than the correlations obtained in Study 1. Similarly, the correlations between the MUSIC variables and cognitive and behavioral engagement were slightly lower than those in Study 1 and varied from 0.13 to 0.57. The scores on the achievement test ranged from 27 to 77 with a mean score of 55.9 ($SD = 8.60$). Achievement was significantly correlated with behavioral engagement, but not cognitive engagement (see **Table 2**).

We analyzed the variables in the path analyses as observed constructs instead of latent constructs because the fit indices were not as good when we used latent constructs in the models. The fit indices were good for model 2a in **Figure 4** that included the MUSIC constructs, cognitive engagement, and achievement (see **Table 3**). The MUSIC constructs explained 31.5% of the variance in cognitive engagement, whereas cognitive engagement only explained 0.8% of the variance in achievement (see **Figure 4**). Success was the only MUSIC construct significantly related to cognitive engagement in the path analysis (see **Figure 4**). The standardized indirect effects for the MUSIC constructs on achievement were also insignificant and the values were as follows: 0.006 for empowerment ($p = 0.286$), 0.001 for usefulness ($p = 0.728$), 0.044 for success ($p = 0.161$), 0.005 for interest ($p = 0.332$), and −0.002 for caring ($p = 0.466$).

The fit indices were good for Model 2b in **Figure 5** that included the MUSIC constructs, behavioral engagement, and achievement (see **Table 3**; Hu and Bentler, 1999; Byrne, 2001; Kline, 2005). The MUSIC constructs explained 31.8% of the variance in behavioral engagement, whereas behavioral engagement explained 2.5% of the variance in achievement (see **Figure 5**). Success and interest were the only two MUSIC constructs significantly related to behavioral engagement in the path analysis (see **Figure 5**). The standardized indirect effects for the MUSIC constructs on achievement were statistically significant for success ($p < 0.01$) and interest ($p < 0.05$), and borderline significant for caring ($p = 0.05$), with the following values: 0.008 for empowerment ($p = 0.370$), 0.010 for usefulness ($p = 0.184$), 0.064 for success ($p = 0.006$), 0.028 for interest ($p = 0.016$), and −0.016 for caring ($p = 0.050$).

TABLE 2 | Correlations among Study 2 variables.

	1	2	3	4	5	6	7
1. Empowerment							
2. Usefulness	0.26						
3. Success	0.50	0.35					
4. Interest	0.49	0.39	0.59				
5. Caring	0.32	0.47	0.22	0.38			
6. Behavioral engagement	0.35	0.40	0.57	0.49	0.20		
7. Cognitive engagement	0.33	0.20	0.56	0.37	0.13*	0.60	
8. Achievement	0.09^{NS}	0.13*	0.24	0.15*	0.09^{NS}	0.18*	0.09^{NS}
M	4.64	5.30	4.27	4.47	5.30	4.42	3.56
SD	0.67	0.71	0.82	0.73	0.52	0.80	0.84
Cronbach's α	0.72	0.77	0.85	0.80	0.75	0.92	0.90

*$p < 0.001$ for all correlations.*Denotes $p < 0.05$. ^{NS}Denotes $p > 0.05$ (not significant).*

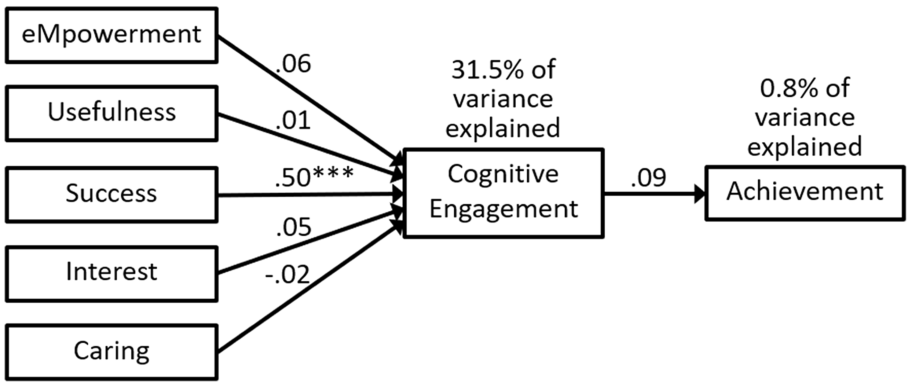


FIGURE 4 | Model 2a: path analysis predicting cognitive engagement and achievement. Statistics are standardized beta coefficients. *** $p < 0.001$.

The fit indices for model 2c in **Figure 6** that included the MUSIC constructs, cognitive and behavioral engagement, and achievement were reasonable for the CFI and SRMR values, but the RMSEA value was a little higher than acceptable (see **Table 3**). The MUSIC constructs explained 31.5% of the variance in cognitive engagement, cognitive engagement explained 35.2% of the variance in behavioral engagement, and behavioral engagement explained 2.5% of the variance in achievement. In the path analysis, success was the only MUSIC construct significantly related to cognitive engagement, cognitive engagement was significantly related to behavioral engagement, and behavioral engagement was significantly related to achievement (see **Figure 6**). The standardized indirect effects for the MUSIC constructs on behavioral engagement were statistically significant only for success ($p < 0.001$), with the following values: 0.038 for empowerment ($p = 0.438$), 0.004 for usefulness ($p = 0.893$), 0.259 for success ($p < 0.001$), 0.031 for interest ($p = 0.546$), and −0.041 for caring ($p = 0.717$). The standardized indirect effects for the MUSIC constructs on achievement were statistically significant only for success ($p < 0.01$), with the following values: 0.006 for empowerment ($p = 0.312$), 0.001 for usefulness ($p = 0.829$), 0.047 for success ($p = 0.007$), 0.005 for interest ($p = 0.392$), and −0.002 for caring ($p = 0.594$).

Discussion

Research Question 1 and 2

The purpose of this study was to investigate the relationships among students' MUSIC perceptions, cognitive engagement, behavioral engagement, and achievement within the context of undergraduate English courses in China. Related to RQ1 (To what extent do students' MUSIC perceptions relate to their cognitive engagement?) and RQ2 (To what extent do students' MUSIC perceptions relate to their behavioral engagement?), the results from Study 2 are similar to those from Study 1 in that all five MUSIC perceptions were significantly correlated with cognitive and behavioral engagement. Furthermore, the MUSIC constructs explained a good amount of the variance in cognitive engagement (31.5%) and behavioral engagement (31.8%). In the path models, success was a significant predictor of cognitive engagement and success and interest were predictors of behavioral engagement. These findings are similar to Study 1 except that interest and empowerment did not predict cognitive

TABLE 3 | Fit indices for the models in **Figures 4**–**6**.

Model	CFI	SRMR	RMSEA [90% CI]	χ^2
Model 2a	0.977	0.050	0.088 [0.039, 0.141]	15.01 ($df = 5$), $p = 0.010$
Model 2b	0.988	0.040	0.064 [0.000, 0.120]	10.29 ($df = 5$), $p = 0.067$
Model 2c	0.938	0.072	0.113 [0.081, 0.147]	47.20 ($df = 11$), $p < 0.001$

CFI, comparative fit index; SRMR, standardized root mean squared residual; RMSEA, root-mean-square error of approximation; and CI, confidence interval.

engagement in Study 2 as they did in Study 1. In sum, we documented that students' perceptions of the motivational climate (as measured by students' MUSIC perceptions) were significantly related to cognitive and behavioral engagement. The success variable was the best predictor of cognitive engagement, and success and interest were the best predictors of behavioral engagement.

Research Question 3

Our third research question asked: To what extent do students' MUSIC perceptions positively predict their cognitive engagement, which then positively predicts their achievement? To answer this question, we first examined the extent to which the data fit the structural model shown in **Figure 4** (Model 2a). The data fit Model 2a reasonably well (see **Table 3**); however, the RMSEA value of 0.088 was a little high and above our pre-identified cutoff value of 0.08. In addition, the path from cognitive engagement to achievement was insignificant, and the indirect effects from all the MUSIC constructs to achievement were insignificant. Given these findings, the answer to RQ3 is that the success variable predicts cognitive engagement, but that cognitive engagement does not predict achievement; and therefore, Model 2a is not a good model.

Research Question 4

Our fourth research question asked: To what extent do students' MUSIC perceptions positively predict their behavioral engagement, which then positively predicts their achievement? **Figure 5** shows the model we tested to answer this question. The data fit Model 2b well (see **Table 3**) with all the fit indices meeting our criteria. Furthermore, the paths from success and interest to behavioral engagement were significant, as was the path from behavioral engagement to achievement. In addition, the indirect effects on achievement were significant for success ($p < 0.01$) and interest ($p < 0.05$) and borderline significant [$p = 0.05$] for caring. Thus, we conclude that Model 2b is a good model, which shows how students' perceptions of success and interest significantly predict behavioral engagement, which then significantly predicts achievement.

Research Question 5

Our fifth research question included both cognitive and behavioral engagement in the same model and asked: To what extent do students' MUSIC perceptions positively predict their cognitive engagement, which then positively predicts their behavioral engagement, which then positively predicts their achievement? Although the CFI and SRMR values were reasonable (see **Table 3**) for the model depicted in **Figure 6** (Model 2C), the value for RMSEA (0.113) was higher than our preidentified cutoff value of 0.08; thus, making the overall fit not very good. As others have noted (e.g., Xia and Yang, 2019), using cutoff values for fit indices is not an exact science. Therefore, although the model fit was not great, it was pretty good. We then examined the significance of the paths between the constructs. The success construct was significantly related to cognitive engagement, which was significantly related to behavioral engagement, which was a significant predictor of achievement. Moreover, there was a significant ($p < 0.001$) indirect effect from success to achievement. Taken together, even though the fit for Model 2c was not as good as it could have been, we documented significant relationships among the constructs as predicted (although only success and not the other MUSIC perceptions were significantly related to cognitive engagement).

Although the path from behavioral engagement to achievement was significant, the magnitude of the standardized beta coefficient ($\beta = 0.16$) was relatively small. It is possible that the measure of achievement was insufficient to capture the *gains* in students' abilities that occurred during the course because it was designed to measure students' abilities at the end of three courses in college English. Because this was only the first course of a three-course sequence of courses, students had not learned all the skills needed to succeed on this test. As evidence, the mean score on the test was only slightly above 50% ($M = 55.9$; $SD = 8.6$) and the scores ranged from 27 to 77 out of a scale that ranged from 0 to 100. Nonetheless, the small standardized beta coefficient is similar in magnitude to those reported in other studies with undergraduates in the United States. For example, Jones et al. (2021) documented a standardized beta coefficient of 0.13 between students' behavioral engagement and achievement. Jones and Carter (2019) reported a slightly higher standardized beta coefficient of 0.24 between behavioral engagement and learning; however, instead of using a measure of achievement, they used a measure of learning that controlled for students' prior knowledge. Therefore, it was likely a better measure of what students learned in the course than the standardized achievement test used in the present study. In sum, compared to other studies, the magnitude of the relationship between behavioral engagement and achievement in the present study is fairly typical and is within the range of 0.00 to 0.30 that Reeve et al. (2020) noted as typical for studies linking engagement to educational outcomes.

LIMITATIONS

The results of this study must be interpreted within the context of its limitations. Although the studies were conducted at two different universities, the results would be more generalizable if more universities and classes were included in the study. Another limitation was that instructors' teaching strategies were

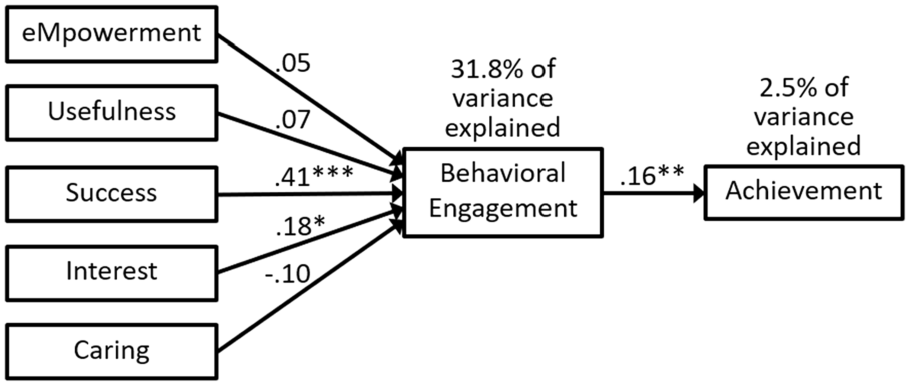


FIGURE 5 | Model 2b: path analysis predicting behavioral engagement and achievement. Statistics are standardized beta coefficients. * $p < 0.05$, ** $p \leq 0.01$, *** $p < 0.001$.

not documented; therefore, we are not able to comment on how instructional practices may have influenced students' perceptions. Some studies have made connections between instructional activities and students' MUSIC perceptions (e.g., McGinley and Jones, 2014; Li et al., 2021; Jones et al., 2022a) and these types of analyses can be helpful to instructors who want to design instruction that motivates students. Finally, in the MUSIC model, external factors such as the culture can influence students' MUSIC perceptions. However, the design of the present studies did not allow us to examine cultural influences on Chinese students' MUSIC perceptions and engagement.

IMPLICATIONS AND CONCLUSIONS

The findings from these studies provide implications for researchers and practitioners interested in effective teaching approaches. We begin this section by discussing some of the theoretical implications for researchers who study the relationships between students' motivational perceptions and engagement. Then, we discuss the implications as they relate to effective teaching.

THEORETICAL IMPLICATIONS

Our findings add to the research studies that have included all the MUSIC constructs in one model to predict engagement. In our studies, success predicted cognitive engagement in both studies and empowerment and interest predicted cognitive engagement in Study 1. In another study with a sample of students very similar to the present studies (i.e., they associated students' MUSIC perceptions with cognitive engagement in an EL class in China), Li et al. (2016) found that success and empowerment predicted cognitive engagement. In a different study with two groups of students, Li et al. (2021) documented that success, empowerment, and interest predicted cognitive engagement with one group of students, and usefulness, interest, and caring predicted cognitive engagement with another group of students who received a different type of instruction. Together, these findings indicate that while success, empowerment, and interest are generally associated with cognitive engagement, these associations can vary somewhat across different EL courses in China. Future studies could include more classes than the present study and be designed to determine whether systematic patterns of relationships between the MUSIC perception variables and engagement exist.

Theoretically, less is known about how the five MUSIC constructs are associated with each other within any one particular course. It is possible that some MUSIC constructs may be antecedents to others in which case increases in any one MUSIC construct could also lead to increases in one or more of the other MUSIC constructs. For example, in the self-determination theory (Ryan and Deci, 2020), constructs similar to empowerment (autonomy), success (competence), and caring (relatedness) are viewed as antecedents to intrinsic motivation, which is often defined similar to how interest was defined in the present study. Therefore, it is possible that empowerment and caring are important in supporting interest in our study. However, the relationships among the MUSIC perceptions are not always straightforward. Researchers studying interest have documented that increasing empowerment through choice can increase interest, but only when certain conditions are met. For example, in one study, choice increased interest, but only when individuals already had a high interest in the task and when the task was perceived as boring (Patall, 2013). Other studies have shown that empowering students through choice can enhance interest when initial success perceptions are high, but not to the same extent when success perceptions are lower (Patall et al., 2014). Perceptions of usefulness have also been found to be related to interest (Patall et al., 2013); yet perhaps only when students have low perceptions of success (Hulleman and Harackiewicz, 2009). Findings from studies such as these demonstrate the complex relationships that can occur between the MUSIC constructs. Although the present studies serve as a proof of concept to demonstrate that relationships exist among these constructs in English courses in China, further

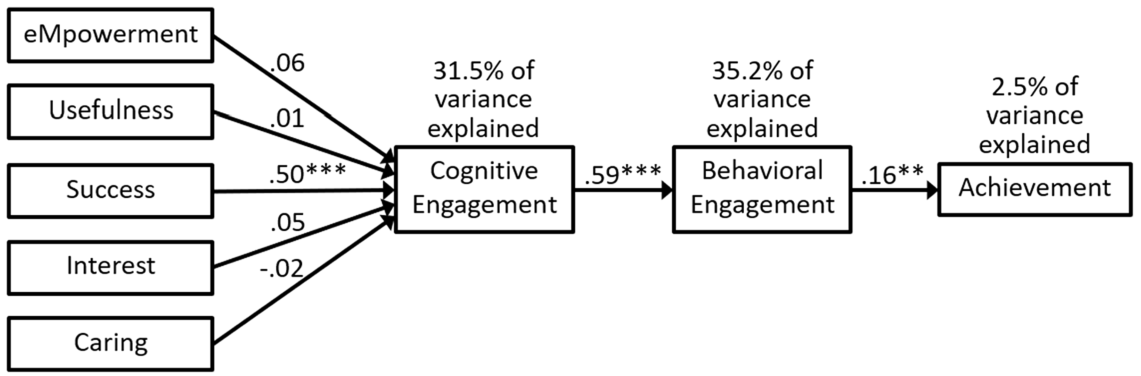


FIGURE 6 | Model 2c: path analysis predicting cognitive and behavioral engagement, and achievement. Statistics are standardized beta coefficients. ** $p \leq 0.01$, *** $p < 0.001$.

studies are needed to systematically examine whether any of the constructs serve as antecedents or whether there are interactions among the constructs.

TEACHING IMPLICATIONS

Because at least some of the MUSIC perceptions are related to engagement in all of these studies, EL instructors may be able to increase students' engagement by implementing strategies that focus on these MUSIC perceptions. Based on the results of these studies, effective instructors could attempt to increase students' perceptions of success, and also, possibly interest and empowerment. Although our findings are based on correlational analyses and do not imply causation, it is reasonable to suspect that EL instructors could increase students' engagement by implementing success, interest, and empowerment strategies. Therefore, an implication is that effective EL instructors engage students by using strategies that increase their perceptions of success, trigger and maintain their interest, and empower them by giving them some choices within course activities and assignments.

For example, strategies that can lead to increases in students' success expectancies include matching the difficulty levels of assignments with the abilities of the students; providing regular, specific feedback to students about their work on assignments; and clearly communicating expectations to students (Jones, 2018, p. 95). There are also many strategies that instructors can use to increase students' situational interest in the course, such as creating activities that pique students' curiosity, using novelty and variety, pacing lessons and lectures appropriately (not too slow or too fast), and limiting lecture time by incorporating more student-centered activities (Jones, 2018). Finally, instructors can empower students by providing them with choices during class and within assignments and incorporating learner-directed approaches (e.g., project-based learning, inquiry approaches). Combinations of these strategies have been shown to increase student engagement in EL courses in China. As an example, when Li et al. (2021) used a novel cell phone technology in class along with a student-centered class activity (students worked in groups to create a summary of what they were learning), students reported higher MUSIC perceptions and engagement than students in a control group that listened to a teacher's lecture. More experimental studies (such as the study by Li et al., 2021) would be helpful to determine how specific instructional strategies can affect students' MUSIC perceptions and engagement.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Virginia Tech IRB. In Study 1, participants completed the survey as part of their normal class activities; and therefore, consent to participate in this study was not required. In Study 2, participants provided their written informed consent to participate in the study.

AUTHOR CONTRIBUTIONS

ML and YG are responsible for the data collection and the first draft. TW is responsible for the statistical analysis. BJ is responsible for the first, second, and third drafts. All authors contributed to the article and approved the submitted version.

FUNDING

The authors would like to thank Shanghai University of Engineering Science and Virginia Tech's Open Access Subvention Fund for providing the financial support needed for the open access publication of this article.

REFERENCES

Ainley, M. (2012). "Students' interest and engagement in classroom activities" in *Handbook of research on student engagement*. eds. A. L. Reschly and C. Wylie (New York: Springer), 283–302.

Al-Hoorie, A. H. (2017). Sixty years of language motivation research: Looking back and looking forward. *SAGE Open* 7:215824401770197–215824401770111. doi: 10.1177/2158244017701976

Al-Hoorie, A. H., and MacIntyre, P. (eds.) (2019). *Contemporary Language Motivation Theory: 60 Years since Gardner and Lambert (1959).* Blue Ridge Summit, PA: Multilingual Matters.

Alrabai, F. (2016). The effects of teachers' in-class motivational intervention on learners' EFL achievement. *Appl. Linguis.* 37, 307–333. doi: 10.1093/applin/amu021

Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: Freeman.

Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*. New York: McGraw-Hill.

Boo, Z., Dörnyei, Z., and Ryan, S. (2015). L2 motivation research 2005-2014: understanding a publication surge and a changing landscape. *System* 55, 145–157. doi: 10.1016/j.system.2015.10.006

Bowlby, J. (1969). *Attachment* (*Vol. I*). New York: Basic.

Byrne, B. M. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming.* Mahwah, NJ: Lawrence Erlbaum.

Chittum, J., and Jones, B. D. (2017). Identifying pre-high school students' science class motivation profiles to increase their science identification and persistence. *J. Educ. Psychol.* 109, 1163–1187. doi: 10.1037/edu0000176

Christenson, S. L., Reschly, A. L., and Wylie, C. (2012a). "Epilogue," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 813–817.

Christenson, S. L., Reschly, A. L., and Wylie, C. (eds.) (2012b). *Handbook of Research on Student Engagement.* New York, NY: Springer.

Cleary, T. J., and Zimmerman, B. J. (2012). "A cyclical self-regulatory account of student engagement: theoretical foundations and applications," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 237–257.

Di Leo, I., Muis, K. R., and Psaradellis, C. (2019). Curiosity… confusion? Frustration! The role and sequencing of emotions during mathematics problem-solving. *Contemp. Educ. Psychol.* 58, 121–137. doi: 10.1016/j.cedpsych.2019.03.001

Dörnyei, Z. (2009). "The L2 motivational self system," in *Motivation, Language Identity and the L2 Self.* eds. Z. Dornyei and E. Ushioda (London, England: Multilingual Matters), 9–42.

Dörnyei, Z., and Ushioda, E. (2011). *Teaching and Researching: Motivation* (*2nd Edn*). London: Pearson.

Eccles, J. S., and Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemp. Educ. Psychol.* 61, 101859–101813. doi: 10.1016/j.cedpsych.2020.101859

Field, A. (2000). *Discovering Statistics Using SPSS for Windows*. London: Thousand Oaks.

Fredricks, J. A., Blumenfeld, P. C., and Paris, A. H. (2004). School engagement: potential of the concept, state of the evidence. *Rev. Educ. Res.* 74, 59–109. doi: 10.3102/00346543074001059

Gardner, R. C. (2019). "The socio-educational model of second language acquisition," in *The Palgrave Handbook of Motivation for Language Learning*. eds. M. Lamb, K. Csizér, A. Henry and S. Ryan (Springer), 21–37.

George, D., and Mallery, P. (2003). *SPSS for Windows Step by Step.* New York, NY: Allyn and Bacon.

Gladman, T., Gallagher, S., and Ali, A. (2020). MUSIC® for medical students: confirming the reliability and validity of a multi-factorial measure of academic motivation for medical education. *Teach. Learn. Med.* 32, 494–507. doi: 10.1080/10401334.2020.1758704

Guo, Y., Xu, J. F., and Xu, X. F. (2020). An investigation into EFL learners' motivational dynamics during a group communicative task: a classroom-based case study. *System* 89, 102214–102215. doi: 10.1016/j.system.2020.102214

Gutiérrez, M., and Tomás, J. M. (2019). The role of perceived autonomy support in predicting university students' academic success mediated by academic self-efficacy and school engagement. *Educ. Psychol.* 39, 729–748. doi: 10.1080/01443410.2019.1566519

Hertling, J. (1996). Chinese students embrace the English language. The Chronicle of Higher Education. Aailable at: http://chronicle.com/article/Chinese-Students-Embrace-the/97026 (Accessed January 3, 2022).

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Hulleman, C. S., and Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science* 326, 1410–1412. doi: 10.1126/science.1177067

Jang, H., Kim, E. J., and Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *J. Educ. Psychol.* 104, 1175–1188. doi: 10.1037/a0028089

Jones, B. D. (2009). Motivating students to engage in learning: The MUSIC model of academic motivation. *Int. J. Teach. Learn. in High. Educ.* 21, 272–285.

Jones, B. D. (2010). An examination of motivation model components in face-to-face and online instruction. *Electron. J. Res. Educ. Psychol.* 8, 915–944.

Jones, B. D. (2012/2021). User guide for assessing the components of the MUSIC Model of Academic Motivation. Available at: http://www.theMUSICmodel.com (Accessed December 15, 2021).

Jones, B. D. (2018). Motivating students by design: practical strategies for professors (2018). CreateSpace. Available at: https://vtechworks.lib.vt.edu/handle/10919/102728 (Accessed April 18, 2022).

Jones, B. D. (2019). Testing the MUSIC model of motivation theory: relationships between students' perceptions, engagement, and overall ratings. *Can. J. Scholarsh. Teach. Learn.* 10, 1–15. doi: 10.5206/cjsotl-rcacea.2019.3.9471

Jones, B. D. (2020a). Engaging second language learners using the MUSIC model of motivation. *Front. Psychol.* 11:1204. doi: 10.3389/fpsyg.2020.01204

Jones, B. D. (2020b). "Motivating and engaging students using educational technologies" in *Handbook of Research in Educational Communications and Technology: Learning Design. 5th Edn.* eds. M. J. Bishop, E. Boling, J. Elen and V. Svihla (Springer), 9–35. doi: 10.1007/978-3-030-36119-8_2

Jones, B. D., Byrnes, M. K., and Jones, M. W. (2019). Validation of the MUSIC model of academic motivation inventory: Evidence for use with veterinary medicine students. Frontiers in veterinary. *Science* 6:11. doi: 10.3389/fvets.2019.00011

Jones, B. D., and Carter, D. (2019). Relationships between students' course perceptions, engagement, and learning. *Soc. Psychol. Educ.* 22, 819–839. doi: 10.1007/s11218-019-09500-x

Jones, B. D., Fenerci-Soysal, H., and Wilkins, J. L. M. (2022a). Measuring the motivational climate in an online course: A case study using an online survey tool to promote data-driven decisions. *Project Leadership & Society* 3:100046. doi: 10.1016/j.plas.2022.100046

Jones, B. D., Krost, K., and Jones, M. W. (2021). Relationships between students' course perceptions, effort, and achievement in an online course. *Comput. Educ. Open* 2:100051. doi: 10.1016/j.caeo.2021.100051

Jones, B. D., Li, M., and Cruz, J. M. (2017). A cross-cultural validation of the MUSIC® model of academic motivation inventory: Evidence from Chinese- and Spanish-speaking university students. *Int. J. Educ. Psychol.* 6, 366–385. doi: 10.17583/ijep.2017.2357

Jones, B. D., Miyazaki, Y., Li, M., and Biscotte, S. (2022b). Motivational climate predicts student evaluations of teaching: Relationships between students' course perceptions, ease of course, and evaluations of teaching. *AERA Open* 8:233285842110731–233285842110717.

Jones, B. D., Osborne, J. W., Paretti, M. C., and Matusovich, H. M. (2014). Relationships among students' perceptions of a first-year engineering design course and their engineering identification, motivational beliefs, course effort, and academic outcomes. *Int. J. Eng. Educ.* 30, 1340–1356.

Jones, B. D., and Sigmon, M. L. (2016). Validation evidence for the elementary school version of the MUSIC® model of academic motivation inventory. *Electron. J. Res. Educ. Psychol.* 14, 155–174.

Jones, B. D., and Skaggs, G. E. (2016). Measuring students' motivation: Validity evidence for the MUSIC model of academic motivation inventory. *Int. J. Teach. Learn. in High. Educ.* 10. doi: 10.20429/ijsotl.2016.100107

Jones, B. D., Tendhar, C., and Paretti, M. C. (2016). The effects of students' course perceptions on their domain identification, motivational beliefs, and goals. *J. Career Dev.* 43, 383–397. doi: 10.1177/0894845315603821

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika* 39, 31–36. doi: 10.1007/BF02291575

King, R. B. (2015). Sense of relatedness boosts engagement, achievement, and well-being: a latent growth model study. *Contemp. Educ. Psychol.* 42, 26–38. doi: 10.1016/j.cedpsych.2015.04.002

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford Press.

Lamb, M. (2019). "Motivational teaching strategies," in *The Palgrave Handbook of Motivation for Language Learning*. eds. M. Lamb, K. Csizér, A. Henry and S. Ryan (Springer), 287–305.

Li, C. (2021). Understanding L2 demotivation among Chinese tertiary EFL learners from an activity theory perspective. *Front. Psychol.* 12:704430. doi: 10.3389/fpsyg.2021.704430

Li, M., Pan, Z., Wang, G., Lang, Y., Ma, D., and Wang, Z. (2021). "The design and evaluation of English class equipped with the iClass platform incorporating motivational strategies." in *ICIMTECH 21: The Sixth International Conference on Information Management and Technology*, August 19; 188, 1–7.

Li, M., Yu, L., Qin, Y., Lu, P., and Zhang, X. (2016). College student academic motivation and engagement in the college English course. *Theory Pract. Lang. Stud.* 6, 1767–1773.

McAuley, E., Duncan, T., and Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Res. Q. Exerc. Sport* 60, 48–58. doi: 10.1080/02701367.1989.10607413

McGinley, J. J., and Jones, B. D. (2014). A brief instructional intervention to increase students' motivation on the first day of class. *Teach. Psychol.* 41, 158–162. doi: 10.1177/0098628314530350

Mohamed, H. E., Soliman, M. H., and Jones, B. D. (2013). A cross-cultural validation of the MUSIC model of academic motivation and its associated inventory among Egyptian university students. *J. Couns. Q. J.* 36, 2–14.

Muenks, K., Wigfield, A., Yang, J. S., and O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *J. Educ. Psychol.* 109, 599–620. doi: 10.1037/edu0000153

Noddings, N. (1992). *The Challenge to Care in Schools: An Alternative Approach to Education*. New York, NY: Teachers College Press.

Pace, A. C., Ham, A.-J. L., Poole, T. M., and Wahaib, K. L. (2016). Validation of the MUSIC® Model of Academic Motivation Inventory for use with student pharmacists. *Curr. Pharm. Teach. Learn.* 8, 589–597. doi: 10.1016/j.cptl.2016.06.001

Patall, E. A. (2013). Constructing motivation through choice, interest, and interestingness. *J. Educ. Psychol.* 105, 522–534. doi: 10.1037/a0030307

Patall, E. A., Dent, A. L., Oyer, M., and Wynn, S. R. (2013). Student autonomy and course value: The unique and cumulative roles of various teacher practices. *Motiv. Emot.* 37, 14–32. doi: 10.1007/s11031-012-9305-6

Patall, E. A., Sylvester, B. J., and Cheon-woo, H. (2014). The role of competence in the effects of choice on motivation. *J. Exp. Soc. Psychol.* 50, 27–44. doi: 10.1016/j.hesp.2013.09.002

Pekrun, R., and Linnenbrink-Garcia, L. (2012). "Academic emotions and student engagement," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 259–282.

Quin, D. (2016). Longitudinal and contextual associations between teacher-student relationships and student engagement: A systematic review. *Rev. Educ. Res.* 87, 345–387. doi: 10.3102/0034654316669434

Reeve, J. (2012). "A self-determination theory perspective on student engagement," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 149–172.

Reeve, J., Cheon, S. H., and Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemp. Educ. Psychol.* 62, 101899–101812. doi: 10.1016/j.cedpsych.2020.101899

Reeve, J., Jang, H., Carrell, D., Jeon, S., and Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motiv. Emot.* 28, 147–169. doi: 10.1023/B:MOEM.0000032312.95499.6f

Renninger, K. A., and Hidi, S. E. (2015). *The Power of Interest for Motivation and Engagement*. New York, NY: Routledge.

Reschly, A. L., and Christenson, S. L. (2012). "Jingle, jangle, and conceptual haziness: evolution and future directions of the engagement construct," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (New York, NY: Springer), 3–19.

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *J. Pers. Soc. Psychol.* 43, 450–461. doi: 10.1037/0022-3514.43.3.450

Ryan, R. M., and Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemp. Educ. Psychol.* 61, 101860–101811. doi: 10.1016/j.cedpsych.2020.101860

Schram, A. B., and Jones, B. D. (2016). A cross-cultural adaptation and validation of the Icelandic version of the MUSIC model of academic motivation inventory. *Icel. J. Educ.* 25, 159–181.

Schraw, G., and Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educ. Psychol. Rev.* 13, 23–52. doi: 10.1023/A:1009004801455

Schunk, D. H., and DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemp. Educ. Psychol.* 60:101832. doi: 10.1016/j.cedpsych.2019.101832

Schunk, D. H., Meece, J. L., and Pintrich, P. R. (2014). *Motivation in Education: Theory, Research, and Applications. 4th Edn.* Columbus, OH: Pearson.

Shell, D. F., and Husman, J. (2008). Control, motivation, affect, and strategic self–regulation in the college classroom: A multidimensional phenomenon. *J. Educ. Psychol.* 100, 443–459. doi: 10.1037/0022-0663.100.2.443

Shell, D. F., Husman, J., Turner, J. E., Cliffel, D. M., Nath, I., and Sweany, N. (2005). The impact of computer-supported collaborative learning communities on high school students' knowledge building, strategic learning, and perceptions of the classroom. *J. Educ. Comput. Res.* 33, 327–349. doi: 10.2190/787L-BCBQ-20FN-FW6C

Skinner, E. A., and Pitzer, J. R. (2012). "Developmental dynamics of student engagement, coping, and everyday resilience," in *Handbook of Research on Student Engagement*. eds. S. L. Christenson, A. L. Reschly and C. Wylie (Boston, MA: Springer), 21–44.

Tabachnick, B. G., and Fidell, L. S. (2007). *Using Multivariate Statistics. 5th Edn.* Columbus, OH: Pearson.

Tao, Y., Meng, Y., Gao, Z., and Yang, X. (2022). Perceived teacher support, student engagement, and academic achievement: A meta-analysis. *Educ. Psychol.* 1–20. doi: 10.1080/01443410.2022.2033168

Tendhar, C., Singh, K., and Jones, B. D. (2017). Using the domain identification model to study major and career decision-making processes. *Eur. J. Eng. Educ.* 43, 235–246. doi: 10.1080/03043797.2017.1329280

Walker, C. O., and Greene, B. A. (2009). The relations between student motivational beliefs and cognitive engagement in high school. *J. Educ. Res.* 102, 463–472. doi: 10.3200/JOER.102.6.463-472

Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educ. Psychol. Rev.* 12, 1–14. doi: 10.1023/A:1009017532121

Wilkins, J. L. M., Jones, B. D., and Rakes, L. (2021). Students' class perceptions and ratings of instruction: Variability across undergraduate mathematics courses. *Front. Psychol.* 12:576282. doi: 10.3389/fpsyg.2021.576282

Wu, P.-S., and Wu, H.-K. (2020). Constructing a model of engagement in scientific inquiry: Investigating relationships between inquiry-related curiosity, dimensions of engagement, and inquiry abilities. *Instr. Sci.* 48, 79–113. doi: 10.1007/s11251-020-09503-8

Xia, Y., and Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behav. Res. Methods* 51, 409–428. doi: 10.3758/s13428-018-1055-2

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Psychology

TYPE Systematic Review
PUBLISHED 22 August 2022
DOI 10.3389/fpsyg.2022.925955

OPEN ACCESS

EDITED BY
Ridwan Maulana,
University of Groningen, Netherlands

REVIEWED BY
Henri Tilga,
University of Tartu, Estonia
Xinpei Xu,
Shanghai Normal University, China

*CORRESPONDENCE
Kai Wang
wangkai.edu@outlook.com

SPECIALTY SECTION
This article was submitted to Educational Psychology, a section of the journal Frontiers in Psychology

RECEIVED 22 April 2022
ACCEPTED 18 July 2022
PUBLISHED 22 August 2022

CITATION
Yang D, Chen P, Wang H, Wang K and Huang R (2022) Teachers' autonomy support and student engagement: A systematic literature review of longitudinal studies.
Front. Psychol. 13:925955.
doi: 10.3389/fpsyg.2022.925955

Teachers' autonomy support and student engagement: A systematic literature review of longitudinal studies

Dong Yang[1], Peng Chen[2], Huanhuan Wang[1], Kai Wang[3]* and Ronghuai Huang[1]

[1]Smart Learning Institute, Beijing Normal University, Beijing, China, [2]College of Education, Capital Normal University, Beijing, China, [3]Center for Teacher Education Research, Beijing Normal University, Beijing, China

Autonomy support is one of the most crucial determinants of teaching practice for student engagement. No literature review on the relations between autonomy support and student engagement existed to the best of our knowledge. Therefore, this study presents a systematic literature review from perspectives of landscapes, methodology characters, patterns of identified studies, and autonomy-supportive strategies. Overall, 31 articles were reviewed. Followed by PRISMA guidelines, the results yielded several interesting facts: First, studies on such topics surged starting from 2015 and were mostly conducted in the United States (32%) and Korea (16%). Publications were scattered but heavily gathered around psychological and educational journals such as the *Journal of Educational Psychology* (9.7%)*; Learning and Instruction (9.7%)*. Most often, studies recruited participants from upper secondary schools (58%). Data were collected using solely questionnaires (93.5%) following a two-wave design (51.6%) and were analyzed by applying structural equation models (48.4%). Moreover, most of the studies failed to provide concrete autonomy-supportive teaching strategies. Instead, quite often studies (93%) investigated its relations with student engagement from a macro perspective. Within mentioned strategies, they were mostly related to the teaching process, there is a limited investigation of autonomy-supportive teaching practice used before and after instruction. This pattern of results suggested an urgent need for more longitudinal studies on specific teaching strategies that hold the potential to maximize student engagement. Limitations and suggestions for future studies were provided accordingly.

KEYWORDS

teaching practice, autonomy support, student engagement, literature review, longitudinal studies

Introduction

Research on student engagement has gained increasing popularity recently as it holds the potential to address problems such as early dropout and poor achievement. The concept of engagement is appealing as it is malleable and sensitive to changes in both teachers' practices (Fredricks et al., 2016). Therefore, engagement was used as a key target for interventions and as an explicit goal of many school improvement programs (Appleton et al., 2008). Recent studies also emphasized the importance of autonomy-supportive teaching practice on student engagement, including teachers' dialogic discourse practice (Böheim et al., 2021) and classroom structure (Cheon et al., 2020). Teacher's autonomy support refers to the degree of latitude teachers give their students during learning activities (Reeve, 2009), including teaching behaviors that detect and nurture kids' needs, interests, and preferences, as well as providing chances in the classroom for students to use their motivations to direct their learning and activities (Reeve et al., 2004). There is evidence that when teachers learn to provide autonomy support in the classroom, it benefits both teachers (e.g., teaching efficacy, teaching skill, and teaching wellbeing) (Rimm-Kaufman and Sawyer, 2004; Cheon et al., 2014) and students (e.g., motivations, classroom engagement, and skill development) (Cheon et al., 2020). More recently, literature review focused on studies applying autonomy-supportive teaching interventions has found that autonomy-supportive teaching is malleable as it can be gained during instruction (Reeve and Cheon, 2021). Participants in most of the intervention studies manifested effective teaching behaviors such as avoiding uttering solutions/answers, being responsive to student-generated questions, spending more time listening, and providing a meaningful rationale (McLachlan and Hagger, 2010; Reeve and Cheon, 2021). Those aforementioned autonomy-supportive teaching behaviors, once learned during the teaching practices, endured (Cheon and Reeve, 2013; Tilga et al., 2020).

One critical factor for boosting student engagement is teaching/motivational styles (i.e., autonomy support from teachers). Teachers who are autonomy-supportive help their students to develop internal motivational resources that promote their engagement in learning (Connell and Wellborn, 1991; Reeve, 2009). Moreover, they present students with meaningful choices between tasks or activities, explain why classroom activities matter, and allow them to pursue their own goals and make decisions on their actions (Reeve et al., 2020).

Although a large body of research has been conducted on the relations between autonomy support and learning engagement, there is still a dearth of studies that synthesize the previous works on such topics. Moreover, recent reviews on student engagement emphasized heavily the technology-mediated environment(e.g., Henrie et al., 2015; Schindler et al., 2017), and it seems that review work on autonomy support is frequently seen in autonomy support in the workplace (e.g., Slemp et al., 2018), and in the field of sports and exercise (e.g., Pérez-González et al., 2019; Raabe et al., 2019). There was one review of the effect of teaching practice on student engagement (i.e., Harbour et al., 2015), that summarized good teaching practices that hold promise to boost student engagement. This particular review, however, failed to provide an intensive picture of how autonomy support could contribute to engagement. A systematic literature review on such a topic provides at least two benefits: first, according to a recent report, the impact of the pandemic on education will last longer than we expected (Dorn et al., 2020), and the damage of the pandemic to individuals goes from learning loss to even loss of earning in students' future working-life (Dorn et al., 2020). Keep students motivated in learning is important during emergency remote teaching, by enhancing autonomy support, teachers can play a role in promoting student engagement; Second, it may provide both researchers and teachers/instructors insights on how to keep students positively engaged in schoolwork, especially in the new teaching normal such as emergency remote teaching (ERT) due to pandemic. In this review, we aim to provide a systematic review of autonomy support and student engagement from four perspectives, namely the landscape of studies, methodologies characters, patterns, and the proposed teachers' autonomy support strategies.

Methods

We conducted a systematic literature review of the published literature on empirical longitudinal studies of autonomy support and student engagement in the past 20 years. A systematic review was chosen because it provides summaries of the state of knowledge in a field from which future research goals may be established, answers issues that individual studies could not; highlights main research flaws that should be addressed in future studies (Page et al., 2021). Therefore, we opt for a systematic review approach to understanding how the studies on teacher autonomy support and student engagement were conducted. To do this, we followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA; Moher et al., 2009) framework when conducting our scoping review. This review covered four topics: (1) the landscape of studies, (2) methodological issues, (3) patterns of previous studies, and (4) the strategies & effectiveness of teachers' autonomy support on student engagement.

Searching strategy

The literature search was performed within databases such as ISI Web of Knowledge, Science Direct, Scopus, and Google Scholar. Those databases were chosen for their breadth in education, psychology, and technology.

TABLE 1 Search terms and strings.

Items	Search terms	Boolean
Autonomy support	"Autonomy-supportive environment" OR "autonomy-supportive interventions" OR "autonomy-supportive teaching" OR "motivational styles" OR "support for autonomy" OR "dialogic discourse practice" OR "supportive instruction practices" OR "classroom climate" OR "teacher support"	AND
Student engagement	"School engagement" OR "engagement in school" OR "student engagement" OR "pupil engagement" OR "learner engagement" OR "emotional engagement" OR "cognitive engagement" OR "behavioral engagement" OR "agentic engagement" OR "academic engagement"	AND
Longitudinal study	"Longitudinal" OR "longitudinal design" OR "longitudinal study" "longitudinal sample" OR "longitudinal associations" OR "longitudinal increase" OR "longitudinal survey" OR "panel study"	

TABLE 2 Inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
Journal articles	Short reports, conference papers, book chapters, etc.
Peer-reviewed	Not peer-reviewed
Empirical studies	Non-empirical studies and theoretical studies
Written in english	Written in other languages
Longitudinal studies	Non-longitudinal studies
Published between 2020 and 2022	Published before 2000 or after the time of writing
Focused on teachers' autonomy support & student engagement	Focused on parents' autonomy support, work engagement, teacher engagement, etc.

We included peer-reviewed journal articles published from January 2000 to March 2022. Three key search terms used on the databases were: "autonomy support" "student engagement" and "longitudinal." Although Similar terms such as "involvement" and "participation" can be found in the literature, we chose to focus only on articles using the word "engagement" in the abstract section, expecting that it would have direct connections with student engagement. We used alternative terms in the searching strings regarding engagement to expand the results, as described in Table 1.

Inclusion and exclusion criteria

To ensure a quality collection of literature, we only chose peer-reviewed journal articles published in English. Since one objective of this study is to explore the topics of teachers' autonomy support and student engagement, we only selected the empirical studies with a longitudinal design. We only chose longitudinal studies as its principal advantage to understand intraindividual change compared with the cross-sectional studies which mainly focused on interindividual differences (Schaie and Hofer, 2001). Detailed inclusion criteria were shown in Table 2.

Screening process

A comprehensive search across databases such as Web of Science, Scopus, ScienceDirect, and additional sources from Google Scholar resulted in 847 articles, and screening of the title and abstracts ($N = 623$) articles resulted in empirical articles that met inclusion criteria. Then we carefully went through each article applying the inclusion and exclusion criteria (see Figure 1), this process yielded a total of 31 articles for final synthesis. The detailed identification flow is shown he Figure 1.

Findings

What are the landscapes of the identified studies?

Countries

We identified countries by the affiliation of the first author. Overall, approximately one out of three studies were conducted in the United States ($N = 10$; 32%); Korea ranks second with an output of five articles (16%). In addition to States and Korea, China and Germany both contributed three articles ($N = 6$; 19%), while the Netherlands added two studies (6%) to the pool. The rest of the studies ($N = 6$; 19%) scattered across Canada, Israel, Peru, Portugal, Spain, and Turkey. No article on such a topic was identified across African countries. See Table 3 for more details.

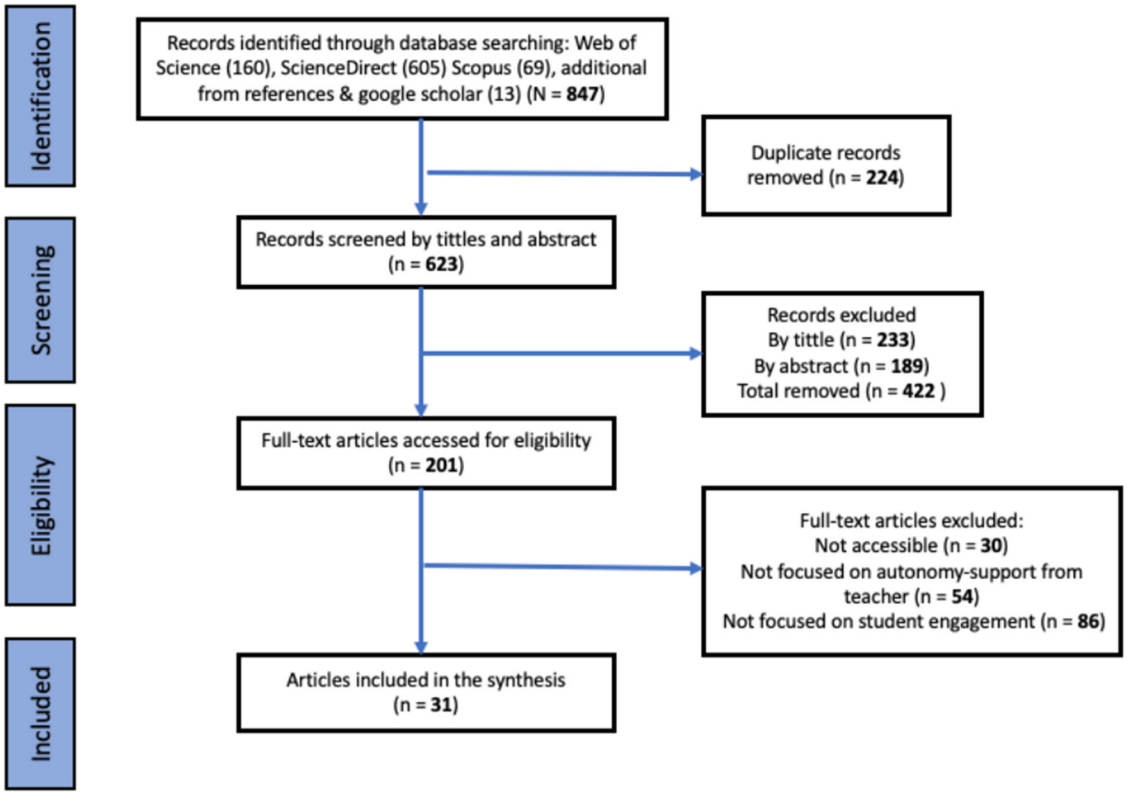


FIGURE 1
PRISMA style article identification flow.

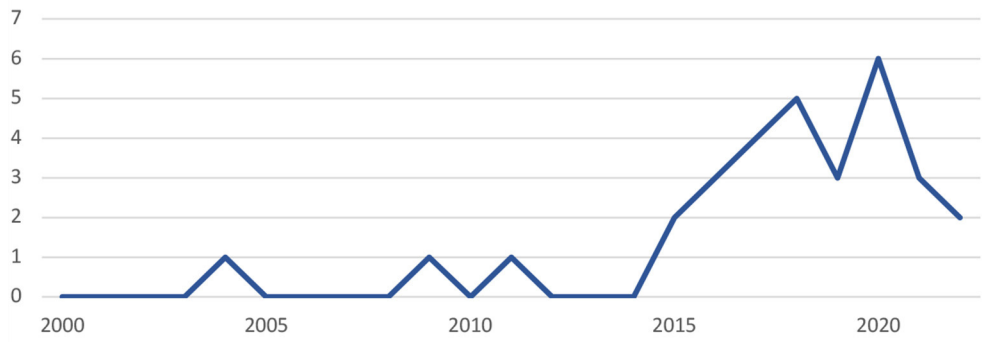


FIGURE 2
Number of publications by year (Till 20 March).

As indicated in Figure 2, at a first sight, the studies on autonomy support and student engagement seem to be scarce in the first decade, with only three articles (10%) screened before the year 2015. Then starting from 2015, academic output on the topic is gaining momentum till the year 2021, which contributed 90% of the total number. Indicating that the topic is getting increasing attention during the past decade. To make the trend clearer, we made a histogram that vividly shows the trend from 2015 to 2022 March in each country. See Figure 3 below.

Journals

In terms of publication, publications are mainly located in Psychological and Educational journals. *Journal of Educational*

TABLE 3 A summary of countries and participants of identified studies ($N = 31$).

Country/ place of study	N	Article	No. of participants: student (teachers)
Canada	1	1. Archambault et al. (2020)	696 (67)
China	3	1. Wei et al. (2020)	1624 (–)
		2. Yu et al. (2015)	356 (–)
		3. Yu et al. (2016)	236 (–)
Germany	3	1. Böheim et al. (2021)	450 (19)
		2. C. Frommelt et al. (2021)	751 (–)
		3. Lazarides and Rubach (2017)	751 (–)
Israel	1	1. Kaplan (2018)	144 (–)
Japan	2	1. Jiang and Tanaka (2022)	199 (87)
		2. Oga-Baldwin and Nakata (2015)	344 (–)
Korea	5	1. Cheon et al. (2016)	1,017 (19)
		2. Cheon et al. (2020)	4,195 (81)
		3. Jang et al. (2012)	500 (–)
		4. Jang et al. (2016)	366 (–)
		5. Reeve et al. (2020)	1,422 (22)
Netherland	2	1. Flunger et al. (2022)	202 (12)
		2. Zee and Koomen (2020)	472 (63)
Peru	1	1. Matos et al. (2018)	336 (–)
Portugal	1	1. Moreira and Lee (2020)	2,676 (–)
Spain	1	1. Núñez and León (2019)	448 (–)
Turkey	1	1. Michou et al. (2021)	257 (–)
United States	10	1. Baker et al. (2017)	120 (6)
		2. Kiefer and Pennington (2017)	209 (–)
		3. Mustafaa et al. (2017)	571 (31)
		4. Patall et al. (2019)	208 (41)
		5. Patall et al. (2018b)	208 (41)
		6. Patall et al. (2018a)	208 (41)
		7. Reeve et al., 2004	–(20)
		8. Ruzek and Schenke (2019)	910 (–)
		9. van Ryzin et al. (2009)	283 (–)
		10. Williams et al. (2018)	113 (3)

"– " means data not reported. One study used only teacher samples (Reeve et al., 2004), others nearly half of the studies (N = 14; 45%) used collected data from the perspectives of both students and teachers.

Psychology (9.7%); *Learning and Instruction* (9.7%) lead the publications, representing three articles separately. Besides, journals such as *Teachers and Teaching: Theory and Practice; Middle Grades Research Journal both* contributed two articles, accounting for 13% of the total sum. The rest of the articles were scattered across different journals, mainly in psychology journals such as *Journal of Adolescence*; *Journal of Youth and Adolescence*. Several publications were identified in the Sports and Science related journals, for example, in *Science Education, Journal of Sport and Exercise Psychology,* etc. Refer the Table 4 below and Table 3 above for full details.

Participants and samples

Identified studies ranged from a small sample size of 20 (Reeve et al., 2004) to large size of 4,195 (Cheon et al., 2020), $M_{sample} = 671$. Fourteen studies (45%) reported both samples from students and teachers, while 16 (52%) studies presented only the data on student participants. One study only used teacher samples (Reeve et al., 2004). The mean age of students is 15.44, mean age for teachers is 37.14 with an average teaching experience of 11.5 years (based on information available). In terms of educational level, more than 90% (N =2 8) of the identified studies were located in the K12 context. Among them, a large body of studies ($N = 58\%$) focused on grade 7–12 level, nearly one out of five ($N = 6$, 19.4%) concerned lower grades 1–6, and four studies (12.9%) used both samples from K6 and K7 to K12 levels. Only three studies focused on the undergraduate groups, representing 9.7% of the total sum. Refer to Table 5 for more info.

What are the features of methodology in current studies?

Study design

In features of methodology, we mainly looked at the study design (waves and period), data type, analysis techniques, theories, and instruments that were utilized in identified studies. First, we found approximately half of the studies ($N = 16$; 51.6%) used a two-wave longitudinal design, one out of four studies ($N = 8$; 25.8%) feature a three-wave design, while four studies applied a four-point measurement (9.7%). Meanwhile, the majority of the studies ($N = 19$; 61.3%) followed a middle-term timespan (that last from several months to a year); seven (22.7%) studies reported a short-term data collection schedule, four (9.7%) with long-term timespan, while only one study features a continuous measurement as described in Table 6.

Data and analysis plan

Second, on data type and analysis plan, studies mostly ($N = 29$, 93.5%) relied on a self-reported questionnaire to capture data, making observation data less appealing comparatively ($N = 2$, 6.5%). Statistically speaking, the studies seem to be obsessed with structural equation models (SEM; $N = 15$, 48.4%), and using multilevel regression analyses such as hierarchical linear modeling (HLM) approaches ($N = 7$; 22.7%). This result is not surprising as a study with repeated measurements usually resulted in a nested data structure (Goldstein et al., 1993). Other analysis techniques, though less favored, were path analysis, HMRA, and repeated measures. Out of the total of 31 articles,

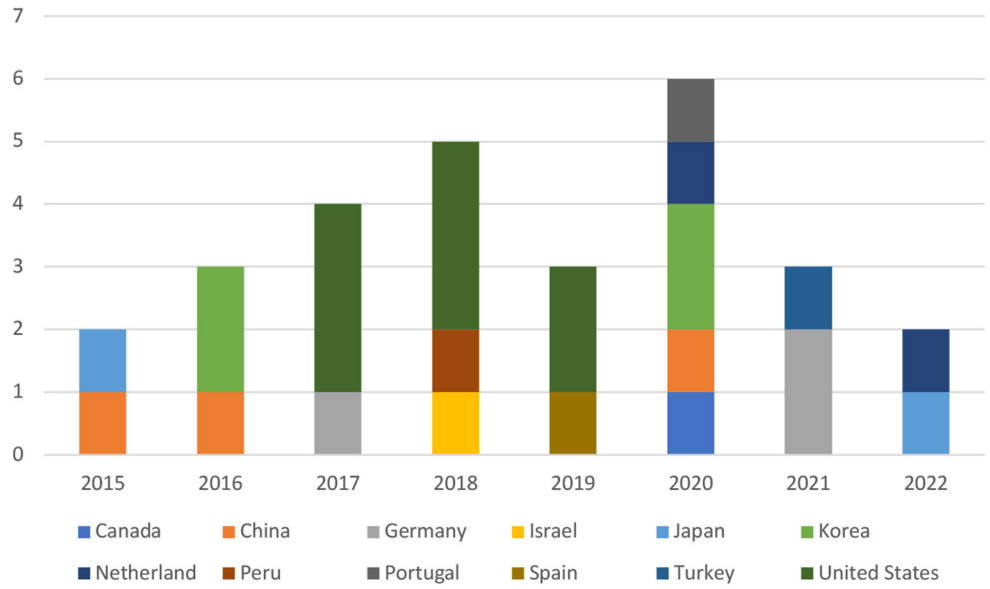


FIGURE 3
Timeline of publication by the geographic area between 2015 and 2022.

26 (83.9%) reported using an intervention/experimental design. See Table 7 below.

Theoretical issues

As expected, more than two out of three (67.7%) studies applied self-determination theory (SDT) as the grounding theory (e.g., Lazarides and Rubach, 2017; Michou et al., 2021; Jiang and Tanaka, 2022), due to its argument that students' motivation and engagement in the classroom are influenced by how they perceive their learning environment and how teachers meet their basic psychological needs (Ryan and Deci, 2000). Besides SDT, three (around 10%) studies referred to social-cognitive theories (Ruzek and Schenke, 2019) or stage–environment fit theory (Yu et al., 2015, 2016) to underpin their studies. In addition, seven articles were unclear on the underpinning theories (e.g., Kiefer and Pennington, 2017; Frommelt et al., 2021). See Table 8 below.

The most frequently investigated aspect of student engagement was behavioral engagement (N = 23, 74.2%), followed by cognitive and emotional engagement (54.8% and 58.1% separately). Almost one out of three studies worked on the agentic perspective, representing 35.5% of the total. Among them, nine (29%) articles measured student engagement from agentic, behavioral, cognitive, and emotional dimensions (e.g., , Cheon et al., 2016; Matos et al., 2018; Núñez and León, 2019), and five studies researched student engagement from the popular "BCE" perspective (i.e., Yu et al., 2015, 2016; Mustafaa et al., 2017; Archambault et al., 2020; Wei et al., 2020); while four studies concerned only on behavioral and emotional aspects (e.g., Reeve et al., 2004; van Ryzin et al., 2009). In addition, four (12.9%) articles were concerned with less frequently used dimensions such as social engagement (e.g., Baker et al., 2017). Details were presented in Table 9 below.

In addition, popular instruments used in studies were presented in Table 10. Due to the space limit, we do not cover this in detail.

Patterns of identified studies

In terms of the pattern of studies, most of the studies feature either bottom-up (i.e., autonomy support impacting student engagement) or top-down (student engagement impacting autonomy support) models that explain the bi-directional relationship between autonomy-supportive teaching strategies and student engagement. For example, a large body of studies (N = 20; 64.5%) utilized structural equation models or path analysis to understand the relations between teacher autonomy support and student engagement that we call the "TS" pattern. A significant amount from the rest of the studies (N = 7; 22.5%) added needs satisfaction into the equations, testing its mediating role in relations between autonomy support and student engagement, this was coded as the "TNS" pattern. Two studies concerned with how the engagement could contribute to teacher autonomy, and in turn, how the perceived autonomy support could boost further engagement. This was named as "ST" pattern. While a significant amount of the rest studies

TABLE 4 Publications of identified articles.

Journals	N of study	Percentage
Contemporary Educational Psychology	1	3.2%
Cyberpsychology, Behavior, and Social Networking	1	3.2%
Educational Psychology	1	3.2%
European Journal of Psychology of Education	1	3.2%
Interdisciplinary Education and Psychology	1	3.2%
International Journal of Behavioral Development	1	3.2%
International Journal of STEM Education	1	3.2%
Japanese Psychological Research	1	3.2%
Journal of Adolescence	1	3.2%
Journal of Applied Developmental Psychology	1	3.2%
Journal of Educational Psychology	3	9.7%
Journal of Experimental Education	1	3.2%
Journal of Sport and Exercise Psychology	1	3.2%
Journal of Research in Childhood Education	1	3.2%
Journal of Youth and Adolescence	1	3.2%
Learning and Individual Differences	1	3.2%
Learning and Instruction	3	9.7%
Learning, Culture, and Social Interaction	1	3.2%
Mathematics Education Research Journal	1	3.2%
Middle Grades Research Journal	2	6.5%
Motivation and Emotion	1	3.2%
Science Education	1	3.2%
Social Psychology of Education	1	3.2%
Teaching and Teacher Education	1	3.2%
Teachers and Teaching: Theory and Practice	2	6.5%

TABLE 5 The level of students studied.

Level of education	Frequency	Percent (%)	Studies (examples)
K6	6	19.4%	Baker et al., 2017; Mustafaa et al., 2017
K7–12*	18	58.0%	Patall et al., 2019; Reeve et al., 2020
K6 & K7–12	4	12.9%	Ruzek and Schenke, 2019; Michou et al., 2021
Undergraduates	3	9.7%	Matos et al., 2018; Jiang and Tanaka, 2022
Total	31	100%	–

(1) * studies reported sample as "middle/high school" were cataloged as K7–K12 level, as corresponded to secondary school.

TABLE 6 Characteristics of study design (waves of data and period).

Waves/types of study	N	Percentage
Interval	4	12.9%
1 interval	16	51.6%
2 intervals	8	25.8%
3 intervals	3	9.7%
Total	31	100%
Continuous	1	3.3%
Short-term	7	22.7%
Middle-term	19	61.3%
Long-term	4	9.7%
Total	31	100%

Interval, waves of measurement unclear; continuous, constant measurement across time; momentary, measured across seconds or minutes; Short-term, measured across days or weeks; middle-term, measured across months to one year; long-term, measured across more than one year.

($N = 7$; 22.5%) added needs satisfaction into the equations, testing its mediating role in relations between autonomy support and student engagement, this was coded as a "TNS" pattern. Unfortunately, two studies (6.5%) failed to indicate any similar pattern (as indicated in Table 11 below).

Which autonomy-supportive strategies were proposed?

Not all studies proposed concrete autonomy-supportive strategies. Still, from the texts, we can summarize several. Grounded in the early work of Reeve et al. (2004) and theories such as SDT (Ryan and Deci, 2000, 2002), most of the strategies used in screened articles include instructional behaviors such as taking the students' perspective (e.g., teaching students' preferred ways), invitational language, provide explanatory rationales, accept mistakes and negative affect, and display patience toward teaching and students (e.g., Jang et al., 2016; Reeve et al., 2020). In addition to autonomy-supportive teaching, a dialogic discourse that is structured, purposeful, interactive, and cumulative as well as guiding was also suggested, for the purpose to maximize student engagement (Böheim et al., 2021). Different from those aforementioned strategies that focused intensively on the teaching process (i.e., classroom teaching), Baker et al. (2017) investigated the effectiveness of teaching framing strategies (e.g., collaborative rule-setting, establishing procedures, or setting goals for interaction and expectations) that occurred before class time, one perspective that deserves more attention.

Most often, the proposed strategies were found effective to promote student engagement (e.g., Baker et al., 2017; Kiefer and Pennington, 2017). However, there are still controversial findings. For example, the study by Ruzek and Schenke (2019) concluded that students' perception of classroom autonomy support was unrelated to students' motivation and engagement among secondary school students, but students' behavioral engagement positively affected the bidirectional connections between their perceptions of autonomy support and academic

TABLE 7 Data type and analysis techniques.

Data type	N	Percentage
Questionnaire data	29	93.5%
Observation data	2	6.5%
Total	31	100%
Analysis techniques		
HLM	7	22.7%
HMRA	2	6.4%
SEM	15	48.4%
Path analysis	1	3.2%
Repeated measures	2	6.4%
Others	4	12.9%
Total	31	100%
Intervention		
Yes	5	16.1%
No	26	83.9%
Total	31	100%

TABLE 8 Theories used in identified studies.

Theory	N of study	Percentage
Self-determination theory	21	67.7%
Social-cognitive theories	1	3.2%
Stage–environment fit theory	2	6.5%
Others (theory not clear)	7	22.6%

TABLE 9 Dimensions of student engagement studies concerned.

Dimensions	Frequency	Percentage
Agentic engagement	11	35.5%
Behavioral engagement	23	74.2%
Cognitive engagement	17	54.8%
Emotional engagement	18	58.1%
Others	4	12.9%

stress. More details are provided below in Table 12. Due to space limitations, we present only examples here.

Discussion

This review explored the basic pillars and landscapes of longitudinal studies on teachers' autonomy support and student engagement. Using a systematic literature review approach, and based on a literature pool of 31 articles (that featuring 20,804 participants), we found the available evidence as presented below:

First, we found that research on the topic mostly occurred in the United States and Korea, the rest of the studies scattered across several European and Asian countries, and there is an underrepresentation of African authors. Although we aimed to search literature from the past two decades, most of the identified studies were conducted from 2014 to 2015, especially on autonomy support and agentic engagement. This is understandable, as the concept of agentic engagement was originally proposed in the year 2013 by Reeve (2013) , thus it is not surprising that the research on such a theme surged since then. In terms of the sample, a large body of the studies recruited upper secondary school students as the samples, and almost every study was set in the classroom environment. This is probably because engagement has been regarded as a concept holding promise for improving reform and significant intervention targets particularly at the secondary level (Appleton et al., 2006; Fredricks et al., 2016). Simultaneously this means that there is a dearth of studies that focus on the underrepresented undergraduate group, and other learning environments such as blended learning and emergency remote teaching (ERT). Thus, we argue that there is a need to shift research focus on autonomy support and student engagement in the context of college teaching and in other new teaching normal such as ERT, a term/field that need consistent attention under the current situation, as students already experienced tremendous learning loss due to the pandemic (Dorn et al., 2020).

Secondly, on characters of methodology. Regards research design, most of the longitudinal studies applied two-wave design across a period of several months to a year. Meanwhile, almost all studies depended on a self-report survey (e.g., questionnaires) for data collection. While questionnaire data is the most common method for assessing student engagement and it is useful in collecting data on students' subjective perceptions, rather than just gathering objective data on behavioral markers such as attendance or assignment completion rates (Appleton et al., 2006). Some argue that questionnaires should only be harassed to access emotional and cognitive engagement which are not directly observable (Fredricks et al., 2016), thus other dimensions such as perceived autonomy support and behavioral engagement are observable sometimes. In addition to the questionnaire, provide observation, semi-structured interview, or even experience sampling methods (ESM; Larson and Csikszentmihalyi, 2014) that capture students' daily experience of teaching practice and engagement (e.g., Patall et al., 2018b), may add extra nuances to our understandings of complex interactions between autonomy support and student engagement. Across studies, the behavioral aspect of engagement was the most investigated (74.2%), probably because it is the only engagement dimension that contributed significantly to school

TABLE 10 Most frequently used instruments.

Dimensions of measurement	Names of instruments (authors)	Study examples
Autonomy support	Learning climate questionnaire (LCQ; Williams and Deci, 1996) ($N = 8$)	Núñez and León, 2019; Cheon et al., 2020
Engagement		
Agentic engagement	Agentic engagement scale (Reeve, 2013) ($N = 11$)	Patall et al., 2018a; Reeve et al., 2020
Behavioral engagement	Engagement vs. disaffection with learning measure (Skinner et al., 2009) ($N = 11$)	Matos et al., 2018; Zee and Koomen, 2020
Cognitive engagement	Metacognitive strategies questionnaire (Wolters, 2004) ($N = 6$)	(Jang et al., 2012)
Emotional engagement	Engagement vs. disaffection with learning measure (Skinner et al., 2009) ($N = 11$)	Cheon et al., 2016; Patall et al., 2018b

Several studies used a self-developed questionnaire or coding frame (N = 6), for example, Oga-Baldwin and Nakata (2015) and Reeve et al. (2004) (coding frame), therefore they are not the scope of discussion of this part.

TABLE 11 Typical patterns of identified studies.

Pattern	N of study	Percentage
Teacher autonomy support –> student engagement (TS)	20	64.5%
Teacher autonomy support –> needs satisfaction–> student engagement (TNS)	7	22.5%
Student engagement –> teacher autonomy support (ST)	2	6.5%
Others (pattern unclear)	2	6.5%

dropout (Archambault et al., 2009), and can be manifested in observable activities.

Moreover, from the patterns of previous studies, we know that most of the studies failed to provide concrete autonomy-supportive teaching strategies, instead, quite often studies followed the schema such as exploring the "*teacher autonomy support –> needs satisfaction–> student engagement*" relations from a general-purpose. Research into how specific teachers' behavior can affect student engagement is becoming increasingly urgent, as the world (e.g., COVID-19 pandemic, district conflicts) and students nowadays are changing, both culturally and psychosocially. Educators need to search for effective ways to meet the challenges presented by the complex world, engage students in the new teaching normal (i.e., emergency remote teaching due to pandemics), and prevent dropouts.

In terms of autonomy support strategies, this review has found several distinct teaching strategies, and mostly they were grounded and underpinned by the early work of Ryan and Deci (2000, 2002) and Reeve et al. (2004). Meanwhile, a large body of aforementioned autonomy-supportive strategies was related to the teaching process, therefore investigation of autonomy-supportive teaching practice undertaken before and after the instruction process was insufficient to some extent. From the literature, one specific strategy proposed is to develop structured and interactive dialogic discourse between teachers and students (Böheim et al., 2021). Early research has repeatedly proven that the quality of a classroom discourse has an impact on students' learning behavior (Mercer and Dawes, 2014; Resnick et al., 2018), and it can be highly effective when students have the opportunity to discuss diverse ways of thinking, elaborating on their perspectives, and develop knowledge constructively and collaboratively (Michaels and O'Connor, 2015; Wilkinson et al., 2017). From this standpoint, further works on how dynamic classroom discourse could contribute to active learning and engagement are pertinent.

Last but not least, our review found that, in general, teachers' autonomy support hold promises to maximize student engagement. In educational practice, this means that teachers are suggested to use effective teaching strategies such as collaborative rule-setting, establishing procedures, or setting goals for interaction and expectations (Baker et al., 2017). When delivering courses, purposeful, interactive, and cumulative, as well as supportive dialogic discourse, is encouraged. Moreover, teachers should bear in mind that when they take the students' perspective (e.g., teaching in students' preferred ways), provide explanatory rationales, accept negative effects, and display patience, they are somehow proving an autonomy-supportive teaching environment. Meanwhile, schools should realize that teachers may have different teaching styles (be it autonomy-supportive or autonomy suppressing), thus carrying out customized teacher training programs is necessary.

Limitations and future work

Several limitations existed in this review study. First, this study used only 31 articles based on inclusion criteria, and we only searched the term *engagement*, instead of using *involvement or participation*, to get additional results. However, it is understandable as the term engagement is more accurate and commonly used across studies (Fredricks et al., 2016). In this study, we only focused on longitudinal studies that

TABLE 12 Autonomy supportive strategies used to promote student engagement (examples).

Authors	Strategies	Effectiveness
Baker et al. (2017)	Teaching framing (i.e., collaborative rule-setting, establishing procedures, or setting goals for interaction and expectations)	Positive
Böheim et al. (2021)	Structured, purposeful, interactive, and cumulative as well as supportive and guiding dialogic discourse	Positive
Cheon et al. (2016), Jang et al. (2016), and Reeve et al. (2020)	Instructional behaviors including: take the students' perspective (e.g., teaching in students' preferred ways), invitational language, providing explanatory rationales, accepting negative effects, displaying patience	Positive
Oga-Baldwin and Nakata (2015), Lazarides and Rubach (2017), and Kiefer and Pennington (2017)	Provide choices, offer respect, show expectations, relevance	Positive
Ruzek and Schenke (2019)	Teachers seek students' perspectives, and respecting their opinions and have standards/expectations for student's efforts, and challenge students to go beyond what they know	Unrelated
van Ryzin et al. (2009)	Teacher-related belongingness (i.e., teacher support)	Positive

feature several waves of data collection, which implies that the same participants are assessed repeatedly and thus other empirical studies were excluded for analysis. Simultaneously, the hierarchical/nested data structure may lead to measuring dependency and thus violates the assumptions underlying the general linear model (GLM; Schnettler et al., 2020). We believe that an independent study/review can compensate for this. In addition, from this view, we learned that a large body of research has focused on autonomy-supportive teaching as a foundation for student motivation and engagement. However, several recent studies have shown that a combination of two teaching styles, namely autonomy-support and structure, can be highly effective for student engagement (Vansteenkiste et al., 2012; Baker et al., 2017; Archambault et al., 2020). In line with this argument and potential constraints of this review, future work might test the effectiveness of structure (the volume and clarity of information provided to students about an activity, including the teacher's expectations concerning educational outcomes and how students are expected to achieve these outcomes, see Jang et al., 2010) with other teaching style factors, or conduct meta-analysis to explore how effective the teaching style (e.g., autonomy support, structure) is on student engagement. In our review, the behavioral aspect of engagement was mostly investigated (74.2%), probably because it is the only engagement dimension that contributed significantly to school dropout (Archambault et al., 2009). Therefore, further investigations are required to testify to the importance of other forms of engagement (e.g., agentic engagement) on student engagement across various learning environments.

Conclusion

Longitudinal studies of teachers' autonomy support and student engagement were explored in a systematic literature review. The main concern of this review is to provide an in-depth review of landscapes, methodology used, trends/patterns of studies, and autonomy-supportive strategies. The main takeaway is that the studies on teachers' autonomy support and student engagement seem to concentrate in countries such as the United States and Korea, while largely underrepresented in African countries. Publications were scattered in the fields of psychology and education. Studies tend to follow a mid-term, two-wave data collection schema using self-reported questionnaires and analyzed by applying SEM. However, most of the studies failed to provide concrete autonomy-supportive teaching strategies, instead, they normally measured autonomy support and student engagement from a broad scale. As stated in the self-determination theory (Ryan and Deci, 2000), student's motivation and engagement in the classroom are influenced by how they perceive their learning environment and how teachers meet their basic psychological needs. In the face of everyday classroom challenges and at times of crisis, students need to display resilience by responding with increased engagement. Therefore, more in-depth exploration of the concrete teaching strategies that boost student engagement, thus preventing school dropout, is becoming increasingly urgent.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Author contributions

DY and KW contributed to the conception and design of the article and interpreting the relevant literature and drafted the manuscript and revised it substantively. PC, HW, and RH contributed to the interpretation of data, and revised it

critically for important intellectual content. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Appleton, J. J., Christenson, S. L., and Furlong, M. J. (2008). Student engagement with school: critical conceptual and methodological issues of the construct. *Psychol. Schools* 45, 369–386. doi: 10.1002/pits.20303

Appleton, J. J., Christenson, S. L., Kim, D., and Reschly, A. L. (2006). Measuring cognitive and psychological engagement: validation of the student engagement instrument. *J. School Psychol.* 44, 427–445. doi: 10.1016/j.jsp.2006.04.002

Archambault, I., Janosz, M., Fallu, J. S., and Pagani, L. S. (2009). Student engagement and its relationship with early high school dropout. *J. Adolescence* 32, 651–670. doi: 10.1016/j.adolescence.2008.06.007

Archambault, I., Pascal, S., Tardif-Grenier, K., Dupéré, V., Janosz, M., Parent, S., et al. (2020). The contribution of teacher structure, involvement, and autonomy support on student engagement in low-income elementary schools. *Teach. Teach. Theory Pract.* 26, 428–445. doi: 10.1080/13540602.2020.1863208

Baker, A. R., Lin, T. J., Chen, J., Paul, N., Anderson, R. C., and Nguyen-Jahiel, K. (2017). Effects of teacher framing on student engagement during collaborative reasoning discussions. *Contemp. Educ. Psychol.* 51, 253–266. doi: 10.1016/j.cedpsych.2017.08.007

Böheim, R., Schnitzler, K., Gröschner, A., Weil, M., Knogler, M., Schindler, A. K., et al. (2021). How changes in teachers' dialogic discourse practice relate to changes in students' activation, motivation and cognitive engagement. *Learn. Cult. Soc. Interact.* 28, 100450. doi: 10.1016/j.lcsi.2020.100450

Cheon, S. H., and Reeve, J. (2013). Do the benefits from autonomy-supportive PE teacher training programs endure?: A one-year follow-up investigation. *Psychol. Sport Exerc.* 14, 508–518. doi: 10.1016/j.psychsport.2013.02.002

Cheon, S. H., Reeve, J., and Song, Y. G. (2016). A teacher-focused intervention to decrease PE students' amotivation by increasing need satisfaction and decreasing need frustration. *J. Sport Exerc. Psychol.* 38, 217–235. doi: 10.1123/jsep.2015-0236

Cheon, S. H., Reeve, J., and Vansteenkiste, M. (2020). When teachers learn how to provide classroom structure in an autonomy-supportive way: benefits to teachers and their students. *Teach. Teach. Educ.* 90, 103004. doi: 10.1016/j.tate.2019.103004

Cheon, S. H., Reeve, J., Yu, T. H., and Jang, H. R. (2014). The teacher benefits from giving autonomy support during physical education instruction. *J. Sport Exerc. Psychol.* 36, 331–346. doi: 10.1123/jsep.2013-0231

Connell, J. P., and Wellborn, J. G. (1991). "Competence, autonomy, and relatedness: a motivational analysis of self-system processes," in *Self Processes and Development* (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.), 43–77.

Dorn, E., Hancock, B., Sarakatsannis, J., and Viruleg, E. (2020). *COVID-19 and Student Learning in the United States: The Hurt Could Last a Lifetime. New Evidence Shows that the Shutdownd Caused by COVID-19 Could Exacerbate Existing Achievement Gaps*. Atlanta, GA: McKinsey and Company.

Flunger, B., Hollmann, L., Hornstra, L., and Murayama, K. (2022). It's more about a lesson than a domain: lesson-specific autonomy support, motivation, and engagement in math and a second language. *Learn. Inst.* 77, 101500. doi: 10.1016/j.learninstruc.2021.101500

Fredricks, J. A., Filsecker, M., and Lawson, M. A. (2016). Student engagement, context, and adjustment: addressing definitional, measurement, and methodological issues. *Learn. Inst.* 43, 1–4. doi: 10.1016/j.learninstruc.2016.02.002

Frommelt, M., Schiefele, U., and Lazarides, R. (2021). Teacher enthusiasm, supportive instructional practices, and student motivation in mathematics classrooms. *Interdiscipl. Educ. Psychol.* 2, 1–5. doi: 10.31532/InterdiscipEducPsychol.2.3.005

Goldstein, H., Bryk, A. S., and Raudenbush, S. W. (1993). Hierarchical linear models: applications and data analysis methods. *J. Am. Stat. Assoc.* 88, 386. doi: 10.2307/2290750

Harbour, K. E., Evanovich, L. L., Sweigart, C. A., and Hughes, L. E. (2015). A brief review of effective teaching practices that maximize student engagement. *Prevent. School Fail.* 59, 5–13. doi: 10.1080/1045988X.2014.919136

Henrie, C. R., Halverson, L. R., and Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: a review. *Comput. Educ.* 90, 36–53. doi: 10.1016/j.compedu.2015.09.005

Jang, H., Kim, E. J., and Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *J. Educ. Psychol.* 104, 1175. doi: 10.1037/a0028089

Jang, H., Kim, E. J., and Reeve, J. (2016). Why students become more engaged or more disengaged during the semester: a self-determination theory dual-process model. *Learn. Inst.* 43, 27–38. doi: 10.1016/j.learninstruc.2016.01.002

Jang, H., Reeve, J., and Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *J. Educ. Psychol.* 102, 588–600. doi: 10.1037.a0019682

Jiang, J., and Tanaka, A. (2022). Autonomy support from support staff in higher education and students' academic engagement and psychological well-being. *Educ. Psychol.* 42, 42–63. doi: 10.1080/01443410.2021.1982866

Kaplan, H. (2018). Teachers' autonomy support, autonomy suppression and conditional negative regard as predictors of optimal learning experience among high-achieving Bedouin students. *Soc. Psychol. Educ.* 21, 223–255. doi: 10.1007/s11218-017-9405-y

Kiefer, S., and Pennington, S. (2017). Associations of teacher autonomy support and structure with young adolescents' motivation, engagement, belonging, and achievement. *Middle Grades Res. J.* 11, 29.

Larson, R., and Csikszentmihalyi, M. (2014). "The experience sampling method BT: flow and the foundations of positive psychology," *in Flow and the Foundations of Positive Psychology* (Dordrecht: Springer). doi: 10.1007/978-94-017-9088-8_2

Lazarides, R., and Rubach, C. (2017). Instructional characteristics in mathematics classrooms: relationships to achievement goal orientation and student engagement. *Math. Educ. Res. J.* 29, 201–217. doi: 10.1007/s13394-017-0196-4

Matos, L., Reeve, J., Herrera, D., and Claux, M. (2018). Students' agentic engagement predicts longitudinal increases in perceived autonomy-supportive teaching: the squeaky wheel gets the grease. *J. Exp. Educ.* 82, 103–120. doi: 10.1080/00220973.2018.1448746

McLachlan, S., and Hagger, M. S. (2010). Effects of an autonomy-supportive intervention on tutor behaviors in a higher education context. *Teach. Teach. Educ.* 26, 1204–1210. doi: 10.1016/j.tate.2010.01.006

Mercer, N., and Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxf. Rev. Educ.* 40, 430–445. doi: 10.1080/03054985.2014.934087

Michaels, S., and O'Connor, C. (2015). "Conceptualizing talk moves as tools: professional development approaches for academically productive discussions," in *Socializing Intelligence Through Academic Talk and Dialogue*, eds L. B. Resnick, C. S. C. Asterhan, and S. N. Clarke (Washington, DC: AERA), 347–361. doi: 10.3102/978-0-935302-43-1_27

Michou, A., Altan, S., Mouratidis, A., Reeve, J., and Malmberg, L. -E. (2021). Week-to-week interplay between teachers' motivating style and students' engagement. *J. Exper. Educ.* 1–20. doi: 10.1080/00220973.2021.1897774. [Epub ahead of print].

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann. Intern. Med.* 151, 264–269. doi: 10.7326/0003-4819-151-4-200908180-00135

Moreira, P. A. S., and Lee, V. E. (2020). School social organization influences adolescents' cognitive engagement with school: the role of school support for learning and of autonomy support. *Learn. Individ. Differ.* 80, 101885. doi: 10.1016/j.lindif.2020.101885

Mustafaa, F., Lozada, F., Channey, J., Skoog-Hoffman, A., and Jagers, R. (2017). Perceptions of teaching practices, teacher support, and school engagement among middle school students: an examination of the developmental designs approach. *Middle Grades Res. J.* 11, 83–98.

Núñez, J. L., and León, J. (2019). Determinants of classroom engagement: a prospective test based on self-determination theory. *Teach. Teach. Theory Pract.* 25, 147–159. doi: 10.1080/13540602.2018.1542297

Oga-Baldwin, W. L. Q., and Nakata, Y. (2015). Structure also supports autonomy: measuring and defining autonomy-supportive teaching in Japanese elementary foreign language classes. *Jpn. Psychol. Res.* 57, 167–179. doi: 10.1111/jpr.12077

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., and Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst. Rev.* 10, 1–11. doi: 10.1186/s13643-021-01626-4

Patall, E. A., Pituch, K. A., Steingut, R. R., Vasquez, A. C., Yates, N., and Kennedy, A. A. U. (2019). Agency and high school science students' motivation, engagement, and classroom support experiences. *J. Appl. Develop. Psychol.* 62, 77–92. doi: 10.1016/j.appdev.2019.01.004

Patall, E. A., Steingut, R. R., Freeman, J. L., Pituch, K. A., and Vasquez, A. C. (2018a). Gender disparities in students' motivational experiences in high school science classrooms. *Sci. Educ.* 102, 951–977. doi: 10.1002/sce.21461

Patall, E. A., Steingut, R. R., Vasquez, A. C., Trimble, S. S., Pituch, K. A., and Freeman, J. L. (2018b). Daily autonomy supporting or thwarting and students' motivation and engagement in the high school science classroom. *J. Educ. Psychol.* 110, 269–288. doi: 10.1037/edu0000214

Pérez-González, A. M., Valero-Valenzuela, A., Moreno-Murcia, J. A., and Sánchez-Alcaraz, B. J. (2019). Systematic review of autonomy support in physical education. *Apunts. Educacion Fisica Deportes* 35, 51–61. doi: 10.5672/apunts.2014-0983.es.(2019/4).138.04

Raabe, J., Schmidt, K., Carl, J., and Höner, O. (2019). The effectiveness of autonomy support interventions with physical education teachers and youth sport coaches: a systematic review. *J. Sport Exerc. Psychol.* 41, 345–355. doi: 10.1123/jsep.2019-0026

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educ. Psychol.* 44, 159–175. doi: 10.1080/00461520903028990

Reeve, J. (2013). How students create motivationally supportive learning environments for themselves: The concept of agentic engagement. *J. Educ. Psychol.* 105, 579–595. doi: 10.1037/a0032690

Reeve, J., and Cheon, S. H. (2021). Autonomy-supportive teaching: its malleability, benefits, and potential to improve educational practice. *Educ. Psychol.* 56, 54–77. doi: 10.1080/00461520.2020.1862657

Reeve, J., Cheon, S. H., and Yu, T. H. (2020). An autonomy-supportive intervention to develop students' resilience by boosting agentic engagement. *Int. J. Behav. Develop.* 44, 325–338. doi: 10.1177/0165025420911103

Reeve, J., Jang, H., Carrell, D., Jeon, S., and Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motiv. Emot.* 28, 147–169. doi: 10.1023/B:MOEM.0000032312.95499.6f

Resnick, L. B., Asterhan, C. S., Clarke, S. N., and Schantz, F. (2018). "Next generation research in dialogic learning," in *The Wiley Handbook of Teaching and Learning*, eds G. E. Hall, L. F. Quinn, and D. M. Gollnick (John Wiley & Sons), 323–338.

Rimm-Kaufman, S. E., and Sawyer, B. E. (2004). Primary-grade teachers' self-efficacy beliefs, attitudes toward teaching, and discipline and teaching practice priorities in relation to the "responsive classroom" approach. *Element School J.* 104, 321–341. doi: 10.1086/499756

Ruzek, E. A., and Schenke, K. (2019). The tenuous link between classroom perceptions and motivation: a within-person longitudinal study. *J. Educ. Psychol.* 111, 903. doi: 10.1037/edu0000323

Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68

Ryan, R. M., and Deci, E. L. (2002). "Overview of self-determination theory: an organismic dialectical perspective," in *Handbook of Self-Determination Research,* eds R. M. Ryan, and E. L. Deci (Rochester, NY: University of Rochester Press), 3–33.

Schaie, K. W., and Hofer, S. M. (2001). "Longitudinal studies in aging research," in *Handbook of the Psychology of Aging, 5th edn*, eds J. E. Birren, and K. W. Schaie (San Diego, CA: Academic Press), 53–77.

Schindler, L. A., Burkholder, G. J., Morad, O. A., and Marsh, C. (2017). Computer-based technology and student engagement: a critical review of the literature: revista de Universidad y Sociedad del Conocimiento. *Int. J. Educ. Technol. Higher Educ.* 14, 1–28. doi: 10.1186/s41239-017-0063-0

Schnettler, T., Bobe, J., Scheunemann, A., Fries, S., and Grunschel, C. (2020). Is it still worth it? Applying expectancy-value theory to investigate the intraindividual motivational process of forming intentions to drop out from university. *Motiv. Emot.* 44, 491–507. doi: 10.1007/s11031-020-09822-w

Skinner, E. A., Kindermann, T. A., and Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educ. Psychol. Measure.* 69, 493–525. doi: 10.1177/0013164408323233

Slemp, G. R., Kern, M. L., Patrick, K. J., and Ryan, R. M. (2018). Leader autonomy support in the workplace: a meta-analytic review. *Motiv. Emot.* 42, 706–724. doi: 10.1007/s11031-018-9698-y

Tilga, H., Kalajas-Tilga, H., Hein, V., Raudsepp, L., and Koka, A. (2020). 15-month follow-up data on the web-based autonomy-supportive intervention program for PE teachers. *Percept. Motor Skills* 127, 5–7. doi: 10.1177/0031512519870914

van Ryzin, M. J., Gravely, A. A., and Roseth, C. J. (2009). Autonomy, belongingness, and engagement in school as contributors to adolescent psychological well-being. *J. Youth Adolescence* 38, 1–12. doi: 10.1007/s10964-007-9257-4

Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., et al. (2012). Identifying configurations of perceived teacher autonomy support and structure: associations with self-regulated learning, motivation and problem behavior. *Learn. Instruct.* 22, 431–439. doi: 10.1016/j.learninstruc.2012.04.002

Wei, D., Zhang, D., He, J., and Bobis, J. (2020). The impact of perceived teachers' autonomy support on students' mathematics achievement: evidences based on latent growth curve modelling. *Eur. J. Psychol. Educ.* 35, 703–725. doi: 10.1007/s10212-019-00437-5

Wilkinson, I. A. G., Reznitskaya, A., Bourdage, K., Oyler, J., Glina, M., Drewry, R., et al. (2017). Toward a more dialogic pedagogy: changing teachers' beliefs and practices through professional development in language arts classrooms. *Lang. Educ.* 31, 65–82. doi: 10.1080/09500782.2016.1230129

Williams, D. R., Brule, H., Kelley, S. S., and Skinner, E. A. (2018). Science in the Learning Gardens (SciLG): a study of students' motivation, achievement, and science identity in low-income middle schools. *Int. J. STEM Educ.* 5, 1–14. doi: 10.1186/s40594-018-0104-9

Williams, G. C., and Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *J. Pers. Soc. Psychol.* 70, 767–779. doi: 10.1037/0022-3514.70.4.767

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientations to predict students' motivation, cognition, and achievement. *J. Educ. Psychol.* 96, 236–250. doi: 10.1037/0022-0663.96.2.236

Yu, C., Li, X., Wang, S., and Zhang, W. (2016). Teacher autonomy support reduces adolescent anxiety and depression: an 18-month longitudinal study. *J. Adolescence* 49, 115–123. doi: 10.1016/j.adolescence.2016.03.001

Yu, C., Li, X., and Zhang, W. (2015). Predicting adolescent problematic online game use from teacher autonomy support, basic psychological needs satisfaction, and school engagement: a 2-year longitudinal study. *Cyberpsychol. Behav. Soc. Netw.* 18, 228–233. doi: 10.1089/cyber.2014.0385

Zee, M., and Koomen, H. (2020). Engaging children in the upper elementary grades: unique contributions of teacher self-efficacy, autonomy support, and student-teacher relationships. *J. Res. Childhood Educ.* 34, 477–495. doi: 10.1080/02568543.2019.1701589

frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 24 July 2023
DOI 10.3389/feduc.2023.1068938

OPEN ACCESS

EDITED BY
Ariel Mariah Lindorff,
University of Oxford, United Kingdom

REVIEWED BY
Grzegorz Szumski,
University of Warsaw, Poland
Kizito Ndihokubwayo,
Iowa State University, United States

*CORRESPONDENCE
Ridwan Maulana
✉ r.maulana@rug.nl

RECEIVED 20 October 2022
ACCEPTED 03 July 2023
PUBLISHED 24 July 2023

CITATION
Maulana R, Kington A, Ko J, Feng X, Helms-Lorenz M, Looker B, Hibbert-Mayne K and Blackmore K (2023) Observing secondary school teachers' effective teaching behavior in the Netherlands, England, and the United States using the ICALT observation instrument. *Front. Educ.* 8:1068938. doi: 10.3389/feduc.2023.1068938

Observing secondary school teachers' effective teaching behavior in the Netherlands, England, and the United States using the ICALT observation instrument

Ridwan Maulana[1]*, Alison Kington[2], James Ko[3], Xiangyuan Feng[1], Michelle Helms-Lorenz[1], Benjamin Looker[2], Kimberley Hibbert-Mayne[2] and Karen Blackmore[2]

[1]Department of Teacher Education, University of Groningen, Groningen, Netherlands, [2]School of Education, University of Worcester, Worcester, United Kingdom, [3]Department of Education Policy and Leadership, The Education University of Hong Kong, Hong Kong, Hong Kong SAR, China

Introduction: The purpose of this study was to examine measurement invariance in observer scoring of effective teaching behavior in three secondary education contexts–the Netherlands, England, and the United States. It also aimed to describe what effective teaching behavior looks like in secondary education across the three education contexts.

Methods: A uniform observation measure called International Comparative Analysis of Learning and Teaching (ICALT) was used to observe teachers.

Results: Results revealed that the hypothesized factor structure of effective teaching behavior was confirmed for the Dutch and English data, but not for the US data. Teachers in the Netherlands showed higher levels of more basic teaching behaviors, but lower levels of more complex teaching behaviors, compared to teachers in England.

Discussion: Implications of the findings are discussed.

KEYWORDS

classroom observation, measurement invariance, effective teaching behavior, secondary education, cross-country comparison

1. Introduction

Research shows that effective teaching behavior plays a central role in student learning and outcomes (e.g., Chapman et al., 2012; Hattie, 2012). International large-scale studies such as the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS) have provided useful insights regarding general trends in educational outcomes in secondary schools around the world (Martin et al., 2016c; Mullis et al., 2016). The PISA and TIMSS studies revealed variations in educational outcomes across education contexts (OECD, 2019). Explanations of differences in educational outcomes across education contexts and countries can be explained by various factors across a number of levels including students', classrooms', schools' and

regional/national characteristics. In general, educational effectiveness research reveals that the classroom level matters the most. Particularly, about 15–25% of differences in student achievement can be explained by the work of teachers (van de Grift et al., 2017). Investigating teaching practices across education contexts may contribute to advance our understanding of variations in educational outcomes. However, little is known regarding differences in effective teaching behavior across education contexts because teaching is inadequately studied in a comparative fashion (Suter, 2019). Comparing teaching practices across contexts offers insights to stimulate cooperation across settings regarding best practices in effective teaching behavior for improved education quality globally (Adamson, 2012; Maulana et al., 2020), and for informing the continuous discourse about teaching and learning across various education contexts (Klette, 2022; Luoto, 2023).

A lack of knowledge regarding differences in effective teaching behavior across education contexts can be explained in several ways. First, the current literature on effective teaching behavior is still rather fragmented and most of the research was conducted in a single education context/country. Second, evaluation of teaching is typically executed using student reports, commonly called Student Evaluation of Teaching (SET, van der Lans et al., 2021), mainly due to the low cost and ease of administration (Maulana and Helms-Lorenz, 2016). Using an observation instrument to measure teaching behavior is unusual (Stroet et al., 2013), partly because it is viewed as costly and highly laborious (Maulana and Helms-Lorenz, 2016), but also because observing classroom teaching behavior is not a common practice around the world due ethics as well as culturally related values. In some contexts, classroom observation is highly valued and widely accepted by schools and teachers, while in others this can be viewed as intrusive and not culturally accepted (Maulana et al., 2023). Nevertheless, observation is seen as a more objective method to describe what is actually happening in the classroom compared to the more commonly used survey approach (Maulana and Helms-Lorenz, 2016).

Various classroom observation instruments exist, but little is known regarding their psychometric properties (i.e., measurement invariance) and potential for international comparisons. To date, there are at least two large-scale classroom observation studies aimed at investigating variations in teaching quality across education contexts. The Teaching and Learning International Survey (TALIS) Video Study, currently known as Global Teaching Insights Video Study, used the Global Teaching Insight (GTI) observation instrument to study teaching in mathematics classrooms across eight education contexts including Chile, Colombia, England (UK), Germany, Japan, Spain, Mexico, and China (OECD, 2020). A further study investigated effective teaching behavior using the Comparative Analysis of Learning and Teaching (ICALT; van de Grift, 2007) in natural classroom settings across school subjects and multiple contexts, including the Netherlands, Indonesia, Republic of Korea, Pakistan, South Africa, and Hong Kong SAR, China (Maulana et al., 2021). These two relatively recent studies offer a promising avenue for cross-country comparison in teaching behavior to stimulate knowledge exchange globally. The present study complements the previous work on effective teaching behavior by including three education contexts, including the Netherlands, England, and the US, and by using a uniform ICALT observation instrument. The aims of the study are twofold: (1) to examine whether the observation measure of teaching behavior can be meaningfully compared (i.e., measurement invariance) in secondary education in the three education contexts; and (2) to investigate differences in effective teaching behavior across the three education contexts.

To our knowledge, this is the first comparative study of teaching behavior using the same observation instrument conducted in these three, specific education contexts. It is our intention that the study will contribute to the discourse regarding the complex nature of teaching by investigating the efficacy of a uniform observation instrument (i.e., ICALT) applied to measure teaching behavior across these contexts.

2. Literature review

In general, there are three common methods for studying classroom teaching, including student surveys, teacher self-reports, and classroom observations (Maulana and Helms-Lorenz, 2016). The three methods have different underlying assumptions and considerations. The current study focuses on classroom observations using a standardized observation instrument.

Comparisons of teaching have been undertaken using a variety of approaches (see Adamson, 2012; Bray et al., 2014). Most straightforward, originating from the early 19th century is the direct comparison of two systems or country specific approaches to determine similarities or differences. An example of this was a comparison of differentiation as a pedagogical strategy to support learners studying for an International Baccalaureate (IB) in a two-center study in Hong Kong SAR, China and Australian schools (Dulfer and Akhlaghi Koopaei, 2021). While this classical approach is still used in educational research (Jortveit et al., 2020; Moberg et al., 2020; Goodwin and Low, 2021), other approaches have since developed. Another is to focus on a particular country as a point of reference, before then comparing other countries to the benchmarks created (Adamson, 2012; Powell, 2020). A third approach sees researchers comparing several countries equally and undertaking complex statistical analysis as a means of examining teaching effectiveness in the different contexts (Adamson, 2012; Powell, 2020). Measurement invariance (e.g., Millsap and Yun-Tein, 2004) is an essential aspect in this approach. With all three approaches it is important to note that research across different countries is only effective if the comparisons add to the field of research; there must be commonality between the compared countries and any international appropriation should be applied sensitively (Adamson, 2012). The current study follows a recent approach by Maulana et al. (2020), by taking the third approach with the ICALT observation instrument.

2.1. Observing effective teaching behavior

Research into effective education has been varied in its approach and areas covered, most notably curricula, teacher behaviors, policy making, leadership and self-efficacy. If effective education can be measured by academic gains and pupil achievement (Coe et al., 2014), then teaching behavior, which has

been well documented as playing a significant role in student learning and outcomes (e.g., Scheerens and Bosker, 1997; Creemers and Kyriakides, 2008; Hattie and Clinton, 2008; Hattie, 2012; Chapman et al., 2016), is a central concept to effective education and teacher effectiveness. Particularly, classroom observations are highly valued in the teacher effectiveness research strand (Muijs et al., 2014; Bell et al., 2019).

It is only since the development of a variety of established observational instruments that the nuances of teacher behavior and quality have been examined. The first classroom observational instruments were developed in the early 1960s, when observational studies examining teacher quality first became popular (van de Grift, 2014). Despite the early instruments lacking validity and reliability, their use and exposure paved a route for the development of more refined and robust methods (Maulana et al., 2020). Since the development of the initial observation measures of teaching, many validated instruments have emerged, situated in strong theoretical and standardized frameworks. These frameworks are generally grounded in the positivist paradigm relying on quantitative approaches. Examples of these include the Protocol for Language Arts Teaching Observation (PLATO; Grossman et al., 2013), the Classroom Assessment Scoring System (CLASS; Pianta et al., 2010), the Framework for Teaching (FfT; Danielson, 2007), the International System for Teacher Observation and Feedback (ISTOF; Muijs et al., 2018), the Global Teaching Insight (GTI Observation System (OECD, 2020)), the Teach Observation System (World Bank, 2022), and the International Comparative Analysis of Learning and Teaching instrument (ICALT; van de Grift, 2014).

Whilst all of these tools differ in structure, theoretical underpinning and implementation techniques, they do share some concepts and characteristics; that is that they are based in the tradition of teaching and teacher effectiveness research and recognized as being measures of effective teaching behavior (Maulana et al., 2014, 2021; van de Grift et al., 2017; OECD, 2020; World Bank, 2022). In addition to more quantitative classroom observation tools, classroom observations based on qualitatively driven frameworks also exist. Examples of these include the Joint Action Framework in Didactics (JAD; Sensevy, 2014), the Cambridge Dialog Analysis Scheme (CDAS; Hennessy et al., 2020), and the ethnographic Documentary Method (DM; Martens and Asbrand, 2022). In the current study, we focus our investigation on the ICALT observation instrument. This framework will be elaborated further below.

2.2. International classroom observation instruments

Observation instruments refer to systematized and standardized tools consisting of a set of predetermined criteria and metric rules for measuring subject-specific and/or generic aspects of teaching skills based on certain views and frameworks of teaching quality. Existing international classroom observation instruments are typically quantitative in nature. There are at least five (high-inference) classroom observation instruments that were designed generically for measuring teaching behavior internationally. These comparative instruments include Virgilio Teacher Behavior Inventory (VITB), International System for Teacher Observation and Feedback (ISTOF), Global Teaching Insights (GTI), Teach, and International Comparative Analysis of Learning and Teaching (ICALT).

An earlier major, large-scale study examining school and teacher effectiveness in Europe, Pacific countries, US, Canada, Australia, and others (Reynolds et al., 2002), suggested that important factors identified through national effectiveness projects could also be found in other national contexts. The study used VITB consisting of three domains: Classroom Management, Instruction, and Classroom Climate (Teddlie et al., 1990). The International School Effectiveness Research Project (ISERP), however, revealed that the VITB had limited external validity when used in only nine countries, each with established school and teacher effectiveness traditions (Reynolds et al., 2002). This finding directed researchers toward further, more detailed examinations of teaching behavior which utilized an observation instrument which was validated across many national contexts. This prompted the subsequent development of ISTOF, used in a study of 20 countries (Reynolds et al., 2002). The instrument has also been adopted in single-country projects. For example, more recently Muijs et al. (2018) reported a satisfactory level of the ISTOF factor structure across a single national site, although they were unable to find ISTOF's 7-factor structure on a consistent basis. Muijs et al. (2018) also highlighted the fact that factor invariance of the ISTOF instrument across countries remains unknown, making it challenging to achieve a cross-country comparison.

Global Teaching Insight was developed by the OECD. The instrument was used to capture an overall picture of teaching quality across eight education contexts including Chile, Colombia, England (UK), Germany, Japan, Spain, Mexico, and China. GTI captures effective teaching behavior in terms of three generic domains including: Classroom Management, Social-emotional Support, and Instruction. The Instruction domain is further divided into three subdomains including Discourse, Quality of Subject Matter, and Student Cognitive Engagement (OECD, 2020). This large-scale study offers general insights into differences in mathematics classroom practices across the eight education contexts. Although construct validity and measurement invariance of the instrument is assumed, we found no specific information regarding the psychometric quality of GTI, particularly on measurement invariance, in more details. Rater quality checks, however, were reported to be conducted systematically (OECD, 2021b).

Teach is a generic classroom observation instrument developed by the World Bank. The instrument was designed to help countries collect data on teaching practices with the aim to improve teaching quality (World Bank, 2022). Teach measures teaching quality in terms of two domains: Time on Task, and Quality of Teaching Practices. The domain Quality of Teaching practices consists of three dimensions: Classroom Culture, Instruction, and Socioemotional Skills. The instrument has mainly been used in low- and middle-income countries, but it was claimed that the instrument can be contextualized for different settings (World Bank, 2022). We also found no specific information on the psychometric quality of the instrument for comparing teaching across education contexts.

International Comparative Analysis of Learning and Teaching is a generic instrument initially developed by four European inspectorates of education including the Netherlands, England,

TABLE 1 Summary of the six ICALT domains.

Domain	Description
Learning climate: safe and stimulating learning climate	Providing a positive classroom environment for learning. This can include promoting respect between all members of the class and the development of students' self-confidence (Cornelius-White, 2007; Hattie and Clinton, 2008; Smith et al., 2008; Ginner Hau et al., 2021)
Classroom management: providing efficient classroom management	The planning and organization of the lesson. This can include quality planning and preparation of lessons, effective time management and pace of lessons, smooth transitions between activities and the swift management of classroom disruption (Wang et al., 1995; Yair, 2000; Marzano, 2003; Opdenakker and Minnaert, 2011; Simonsen et al., 2020)
Clarity of instruction: clear and structured instruction	This relates to the quality of instruction given by the teacher. It can include a clear and organized structure to the lesson, clear explanations of subject content and the effective communication of group and individual classroom tasks (Kindsvatter et al., 1988; Mortimore et al., 1988; Rosenshine, 2010, 2012) alongside the regular assessment of student understanding (Hattie and Clinton, 2008; Smith et al., 2008)
Activating teaching: intensive and activating teaching	This includes behavior for learning principles (Maulana et al., 2017) which ensure the learning of students is maximized, cognitive load is minimized (Bolkan, 2016) and prior knowledge is used effectively to develop schemas for learning (Paas and van Merriënboer, 2020)
Differentiated instruction: adjusting instructions and student processing to inter-student differences	This is an inclusive aspect of teaching behavior, where the teacher involves students from a diverse range of personalities and backgrounds in the lesson. It involves adapting their teaching approaches and learning episodes to suit students' diverse range of learning needs in order to facilitate knowledge acquisition and improved student outcomes (Tomlinson et al., 2003; Reis et al., 2011; Ismajli and Imami-Morina, 2018)
Teaching learning strategies: teaching students learning strategies	This domain focuses on metacognitive learning and teaching strategies employed by the teacher. By employing these strategies and providing students with learning opportunities, it has been shown to provide students with the scaffolded support they need to succeed academically (Rosenshine and Stevens, 1986; Houtveen and van de Grift, 2007; Bae and Kwon, 2021)

Flanders, Lower Saxony (Germany) (van de Grift, 2007). The first ICALT version was developed to capture mathematics teaching quality in primary schools. The instrument was developed further for use in secondary schools across subjects (van de Grift et al., 2014; Maulana et al., 2017). The ICALT observation instrument has been found to be sufficiently invariant across several countries in both primary and secondary education contexts (van de Grift, 2014; van de Grift et al., 2017). Maulana et al. (2020) has since used the ICALT observation instrument with notable success and limitations. They discovered that observers rated Republic of Korea the highest of four countries (compared with Netherlands, South Africa, and Indonesia) in terms of effective teaching behavior. They also found that Differentiated Instruction was observed as being low in both Republic of Korea and in the Netherlands, though observers rated Republic of Korea higher than the Netherlands within this domain. While measurement invariance for two further countries involved in the study (Hong Kong SAR, China and Pakistan) was not found when retaining the full sets of ICALT items, the measurement invariance found for four of the six partaking countries provides further interest in using the ICALT observation instrument across multiple national contexts.

2.3. Effective teaching behavior: the ICALT framework

The ICALT framework is grounded in the evidence-based teacher effectiveness research (TER). Based on reviews of TER, six observable effective teaching behavior domains were synthesized (van de Grift, 2007). These six domains, discussed in depth by Maulana et al. (2021), are included in the ICALT observation tool. The six ICALT domains resonate with findings from other studies of effective teaching behavior (e.g., Klieme et al., 2009; Pianta and Hamre, 2009; Danielson, 2013; Ko and Sammons, 2013; Kington et al., 2014; Muijs et al., 2018). A summary of each of the six domains is given in **Table 1** below.

van de Grift et al. (2014) linked the theory of teachers' concerns (Fuller, 1970) with the ICALT theoretical framework. Fuller's theory posits that teachers' concerns may develop from self-related (focused on self), to task-related (focused on task), and then to impact-related (focused on impact for students) concerns. By applying the Rasch modeling approach, it has been shown that the six domains of effective teaching behaviors can be separated into two different levels of difficulty (van de Grift et al., 2014). The Rasch modeling offers a unique possibility for arranging teaching skill scores on a single dimension and for estimating individuals' skill levels on a latent variable. Hence, this modeling approach makes it possible to link the teacher's concern theory with teaching behavior domains. The first three (Learning Climate, Classroom Management, Clarity of Instruction) have been identified as lower levels of teaching behavior difficulty. This contrasts with the final three domains (Activating Teaching, Differentiated Instruction, Teaching Learning Strategies) providing greater levels of difficulty for teachers (van de Grift et al., 2014; van der Lans et al., 2018).

The ICALT observation instrument has been used to measure teaching behavior for nearly 20 years. In the two decades since its conception, the tool has been developed and used to compare teaching behavior across primary schools in Europe (van de Grift, 2007) and subsequently validated using confirmatory factor analysis (CFA) (van de Grift, 2014). The tool has also been validated for secondary education in the Netherlands (Maulana et al., 2017), Republic of Korea (van de Grift et al., 2017) and Indonesia (Irnidayanti et al., 2020), and for university education in the Netherlands (Noben et al., 2021). In a similar study to the one reported here, the ICALT observation measure was used

to measure invariance of observed secondary teaching behavior across the Netherlands, Republic of Korea, South Africa, Indonesia, Hong Kong SAR, China, and Pakistan (Maulana et al., 2021). Four of the six countries showed measurement invariance, suggesting that the ICALT instrument might be an effective tool for cross-national comparisons of teaching behavior more globally.

3. The current study

3.1. Research questions

Given the gap in comparative educational research for a cross-nationally validated observation instrument to measure teaching behavior, this paper reports on results from an international study that employed the ICALT instrument three across different national contexts. The research questions which guided the analysis were:

1. To what extent is there evidence of an invariant internal structure regarding effective teaching behavior in the Netherlands, England, and the United States?
2. What are similarities and differences regarding the six domains of effective teaching behavior across the three education contexts?

3.2. Contexts of the study

3.2.1. The Netherlands

International comparisons in secondary education indicated that students attending Dutch schools perform above average, comparable to other high performing European and Asian educational systems (Martin et al., 2016a,b; Mullis et al., 2016; OECD, 2016). The majority of students in their teenage years achieve the basic skills in reading, mathematics and science (scores 485, 519, and 503, respectively) (OECD, 2019). The Dutch educational system is highly tracked (i.e., students are split by ability in a large number of different educational tracks from the age of twelve), does not apply a national curriculum, shares national educational standards and gives extensive autonomy to schools and teachers (OECD, 2014, 2016). The high level of decentralization is balanced by a strong school inspection mechanism and a national examination system. The teaching profession does not have a high status in the general public opinion (Brouwer et al., 2016). Nevertheless, the quality of teachers is generally high with the large majority mastering the basic teaching skills to a good standard (OECD, 2016; Inspectie van het Onderwijs, 2018). For secondary school practitioners, there are two types of teaching qualification; first degree qualification (highest degree, license for teaching in all grades), and second degree qualification (license for teaching in lower grades) (Brouwer et al., 2016). First degree qualified teachers generally show higher levels of effective teaching behavior compared to second degree teachers (Helms-Lorenz et al., 2020). It is also possible to teach in secondary education without teaching qualification as a lateral second-career entrant (in Dutch: zij-instromer) or a guest teacher (Rijksoverheid, 2023) with the requirement to qualify within 2 years.

3.2.2. England

In England, secondary school education (pupils aged 11–18 years) takes place in either a state-funded or independent school, with the majority of pupils (93.6%) being educated in state schools (Gov.uk, 2021).[1] Schools maintained by a local authority follow the national curriculum which is set by the Department for Education (DfE). Within the state school system, over half (57.7%) have been designated with "free school" or "academy school" status (see text footnote 1 Gov.uk, 2021) and have more flexibility to deviate from the National Curriculum. Examination boards must set their learning specifications against the DfE's subject content, which means that students must cover the same subject content for the General Certificates of Secondary Education (GCSE) qualifications. Furthermore, the Office for Standards in Education (Ofsted) carry out a rigorous inspection program of all state-funded schools, ensuring the curriculum is broad and that teachers have "good knowledge of the subject" (Ofsted, 2021, p. 40) they teach. Academic outcomes of pupils in the UK have seen an improvement, with reading ability (score 504) rising from 25th to 14th place amongst OECD countries (OECD, 2021a), and scores for pupils in England being the highest of the UK nations (math and science scores 502 and 505, respectively) (OECD, 2019). In mathematics and science, the UK is above the average for OECD countries, and is showing a continuing upward trend (OECD, 2021a). All teachers in England need to have an undergraduate degree to teach. Although it is possible to be employed as an unqualified teacher, this is not common practice as most teachers must undertake formal training and be awarded Qualified Teacher Status (QTS). There is strict performativity and accountability agenda in schools, which is overseen by the Ofsted inspectorate. This is paired with an increasing rate of attrition; for example, in 1996, 9% of teachers left the profession after 1 year. This has increased to 16% in 2019 (Gov.uk, 2021).

3.2.3. The United States

Education is more decentralized in the United States than in most European countries. State-funded public schools represent a high percentage, about 87%, of K-12 education, while tuition and foundation-funded private schools account for about 10%, and home-schooling is roughly 3% (U.S. Department of Education, 2013). Federal spending on education is relatively modest at about one-sixth of the state spending. Therefore, it is not surprising that each state sets its own State Compulsory School Attendance Laws (U.S. Department of Education, 2013). Compulsory education starts between five and eight and ends between sixteen and eighteen. Each state can set its curriculum and staffing policies, except private schools are free to determine theirs, which can also obtain accreditation through independent regional accreditation authorities. While the U. S. spent more per student on education than any other country, 15-year-old American students ranked the 31st in the world in reading literacy, mathematics, and science in The Program for International Student Assessment (OECD, 2018). The average American students scored 487.70, compared with the OECD average of 493 in overall knowledge and skills. Teacher quality is found to significantly improve student achievement (Rockoff, 2004; Hanushek and Rivkin, 2010; Chetty et al., 2014;

1 www.Gov.uk

TABLE 2 Demographic information of participating teachers.

Countries	N_{school}	School denom		N_{teache}	Teacher gender		Teacher subject		Teacher experience		Class size
		Public	Private		Male	Female	Science[a]	Non-science[b]	Inexperienced	Experienced	
Netherlands	99	99	0	200	96 (48.0%)	104 (52.0%)	74 (37.0%)	126 (63.0%)	100 (50.0%)	100 (50.0%)	22.09 (5.30)
England	15	15	0	115	50 (43.5%)	65 (56.5%)	61 (53.0%)	54 (47.0%)	6 (5.2%)	109 (94.8%)	23.18 (3.05)
United States	64	*NA*	*NA*	105	*NA*	*NA*	50 (47.6%)	55 (52.4%)	*NA*	*NA*	*NA*
Total	178			420							

[a]Natural science subjects normally go to the category of *beta* in the Dutch context.
[b]Non-science subjects can be further categorized into alpha subjects (i.e., humanities) and gamma subjects (social sciences) in the Dutch context.

Bruns and Luque, 2015). While some states have improved their teacher preparation programs (e.g., Commission of Teacher Credentialing of California, 2021), an earlier study by Harris and Sass (2011) failed to establish a consistent relationship between formal professional development training and teacher productivity. Although teachers reported being largely satisfied with their jobs and career, only a minority believed that teaching is valued by U.S society (OECD, 2013). The current state of the teaching profession in the U.S. is near its lowest levels in over a half century (Kraft and Lyon, 2022).

4. Materials and methods

The current study is part of a larger longitudinal project on effective teaching behavior across countries called ICALT3/Differentiation.[2] In this project, effective teaching behavior was studied from researcher observations, teacher perceptions, and student perceptions. The current study focuses on reporting the observation data of one particular measurement moment (cross-sectional) from three educational contexts: the Netherlands, England and the US. Particularly, data collected during the second year of fieldwork was used in all three countries.

4.1. Sample and procedure

The current study included 420 teachers from 178 secondary schools in the Netherlands ($N_{teacher}$ = 200), England ($N_{teacher}$ = 115), and the United States ($N_{teacher}$ = 105). In the Netherlands, the cross-sectional data were collected from different cohorts between 2015 and 2018 school years. A random sample of 200 observed teachers from a total of 2157 teachers was selected to compensate for sample imbalance with the other two datasets. The observed teachers taught in schools across 12 provinces in the country. In England, the data were also collected from different cohorts between 2017 and 2019 school years. The observed teachers were from schools across the West Midlands region. Hence, the data is not a nationally representative sample as the other eight regions (including London) were not included. In the US, the video data were collected in 2011 as part of the Measurement of Teaching (MET) project.[3] A random selection of 105 teachers from 12 US regions was done from a pool of available video data.

Table 2 illustrates the denomination of schools, the gender, subject areas and teaching experience of teachers, as well as the average size of their classes. The dividing line between inexperienced and experienced teachers is set at 5 years of teaching experience. The background information of schools and teachers, except teachers' subject expertise, was missing in the US dataset. All participating schools in the Netherlands and England were publically funded schools. The distribution of teachers is relatively balanced across gender, subjects and experience levels, except for teacher subject in the Netherlands and teacher experience in the US. Specifically, non-science teachers are overrepresented in the Dutch

2 https://www.rug.nl/gmw/lerarenopleiding/onderzoek/psychometrisch/

3 https://usprogram.gatesfoundation.org/news-and-insights/usp-resource-center/resources/met-project-data

data, while experienced teachers are overrepresented in English data.

In the Netherlands and England, teachers were invited to participate in the study in accordance with each institution's approved ethical procedures. This included sending information to teachers and having them sign a consent form that served as an agreement between the researchers, the teacher and the teacher's school. Consent was given voluntarily and with full knowledge. In the Netherlands and England, lessons were observed in real-time in their natural classroom settings.

For the US, we cooperated with the MET project coordinator and researcher. The MET project collected video data on classroom teaching. The trained observers conducted a secondary analysis and coded the selected lesson videos/teachers of the 10000+ lessons originally rated with CLASS (Pianta et al., 2010).

4.2. Measures

Effective teaching behaviors were observed using the International Comparative Analysis of Learning and Teaching (ICALT) observation instrument (van de Grift et al., 2014). The instrument includes 120 low inferential observable teaching indicators and 32 high inferential observable teaching behaviors. These high inference items were rated on a four-point scale ranging from 1 ("mostly weak") to 4 ("mostly strong,") and represent the six aforementioned domains of teaching behavior: Learning Climate (four items), Classroom Management (four items), Clarity of Instruction (seven items), Activating Teaching (seven items), Differentiated Instruction (four items), and Teaching Learning Strategies (six items). The six-factor structure of observed teaching behavior has been validated by prior research in other national contexts (Maulana et al., 2020).

The ICALT instrument was used for the current study for several reasons (see Maulana et al., 2021). In summary, this observation tool is, relatively speaking, simple to use in the classroom context, has been translated into many languages, and has already been validated through its use in a number of previous international studies (Maulana et al., 2021, 2022). The instrument was developed with a strong grounding in evidence-based teacher effectiveness research, and its validity has been demonstrated in both primary and secondary education settings, adding to its external validity (van de Grift, 2014; Maulana et al., 2017, 2021). The ICALT tool has also been shown to be appropriate for use in both comparative educational cultures—for example, in several countries in Europe (see van de Grift et al., 2014)–and in contrasting cultures, such as Indonesia, Republic of Korea and South Africa, amongst others (Maulana et al., 2021). Furthermore, the instrument has been established as valuable tool for use in both research and practice in the Netherlands and in Republic of Korea (Maulana et al., 2020). It is frequently used as a diagnostic measure for the professional development of teachers and pre-service teachers (Maulana et al., 2017; Helms-Lorenz et al., 2019), highlighting its versatility and the relative ease with which observers can be trained.

4.3. Observer training

Before using the ICALT observation instrument, each observer underwent onsite training led by two expert trainers. The training was conducted in accordance with the same standards, structure and procedure in all three countries. In order to be qualified for training, observers had to meet two criteria. First, they needed to demonstrate sufficient knowledge of effective teaching. Second, the observers needed at least 3 years of teaching experience, preferably in secondary education. Compared to the Netherlands and England, observers in the US were less experienced teachers but highly knowledgeable with regard to teaching behavior research. In England, all observations were conducted by the research team; no observations were carried out by teachers.

Training consisted of three phases: planning, execution, and evaluation. Phase 1 was dedicated to exploration of the theoretical basis and context of the instrument to deepen trainees' understanding of the theory underlying the ICALT instrument. Phase 2 was devoted to administration of the instrument, including how to rate indicators of teaching behavior using the applied scoring rules. More specifically, trainees coded the teacher behavior in two videotaped lessons using the observation instrument. When analyzing the observation scores of trainees, a consensus

TABLE 3 Categorical confirmatory factor analysis for three countries.

	Model	*N*	χ^2 (df)	RMSEA with 90% CI	SRMR	CFI	TLI
Netherlands	All items	200	911.259* (449)	0.072 [0.065, 0.078]	0.085	0.925	0.917
England	All items	115	588.954* (449)	0.052 [0.040, 0.063]	0.098	0.962	0.958
	11 correlations set to 1	115	599.547* (460)	0.051 [0.039, 0.063]	0.098	0.962	0.959
United States	All items	105	1303.847* (449)	*0.135 [0.126, 0.143]*	*0.232*	*0.594*	*0.552*
	Items 1, 14, 18 removed	105	821.182* (362)	*0.110 [0.100, 0.120]*	*0.188*	*0.705*	*0.669*
	Items 1. 14. 18, 20, 25, 32 removed, CLAR with ORG set to 1	105	569.205* (284)	*0.098 [0.086, 0.109]*	*0.171*	*0.778*	*0.746*
	Items 1, 14, 18, 20, 25, 32 removed, v17 with v27, v24 with v29, CLAR with ORG set to 1	105	507.825* (283)	*0.087 [0.075, 0.099]*	*0.157*	*0.825*	*0.799*
	Items 1, 14, 16, 18, 19, 20, 24, 25, 32 removed	105	412.907* (215)	*0.094 [0.080, 0.107]*	*0.158*	*0.801*	*0.766*

Unacceptable fit indices are in *italics*. *$p < 0.05$.

TABLE 4 Standardized factor loadings of separate CFA for the Netherlands and England.

Country and domains	Standardized factor loadings							Domain correlations					Variance explained
								1	2	3	4	5	
Netherlands (N_{total} = 200)													
1. Learning climate (four items)	0.707	0.816	0.759	0.799									59.5%
2. Classroom management (four items)	0.774	0.843	0.683	0.756				0.814					58.7%
3. Clarity of instruction (seven items)	0.649	0.745	0.657	0.802	0.732	0.793	0.699	0.850	0.887				52.9%
4. Activating teaching (seven items)	0.694	0.688	0.810	0.804	0.769	0.727	0.542	0.797	0.705	0.874			52.4%
5. Differentiated instruction (four items)	0.736	0.862	0.847	0.799				0.371	0.399	0.426	0.610		66.0%
6. Teaching learning strategies (six items)	0.805	0.853	0.752	0.817	0.734	0.847		0.444	0.368	0.467	0.735	0.611	64.4%
England (N_{total} = 115)													
1. Learning climate (four items)	0.622	0.775	0.829	0.594									50.7%
2. Classroom management (four items)	0.694	0.511	0.751	0.637				*1.000*					42.8%
3. Clarity of instruction (seven items)	0.649	0.590	0.692	0.648	0.651	0.723	0.648	*1.000*	*1.000*				43.4%
4. Activating teaching (seven items)	0.704	0.661	0.560	0.612	0.812	0.831	0.621	0.953	*1.000*	*1.000*			47.9%
5. Differentiated instruction (four items)	0.600	0.793	0.590	0.630				0.959	*1.000*	0.980	*1.000*		43.3%
6. Teaching learning strategies (six items)	0.651	0.768	0.557	0.769	0.637	0.525		*1.000*	*1.000*	*1.000*	*1.000*	0.977	43.3%

All factor loadings and factor correlations reported in the table are significant (p <0.001) except the ones in *italics*.

10.3389/feduc.2023.1068938

level of 70% within the group and between the group and the expert norm was set as a sufficient cut-off. Discussions to resolve significant differences and improve the consensus level were conducted subsequently. Finally, the evaluation phase involved the investigation of rating patterns and significant deviations from the average pattern. A small number of observers who deviated from the average were followed up and extra guidance was given to this group prior to conducting the observation in natural classroom settings. Observers failing to meet the minimum consensus of 70% were not invited to conduct observations. The consensus levels between the trainees and the expert norm are 86% in the Netherlands, 71% in England, and 75% in the US, respectively.

4.4. Analytic approach

The measurement model of the ICALT instrument was first subjected to cross-country validation using categorical confirmatory factor analysis (CFA). After confirming the validity of the measurement model in each country data, this study tested the measurement invariance by performing multi-group confirmatory factor analysis (MGCFA) on the combined country data using MPlus version 8 (Muthén and Muthén, 2017). We followed the works of Millsap and Yun-Tein (2004) and Xing and Hall (2015) as references when estimating the models. Three levels of measurement invariance (configural, metric, and scalar) were tested successively. The test of configural equivalence examined whether the same factor structure was applied across countries (i.e., same factorial structure). Metric invariance test verified whether teaching behavior from different countries was rated by the items in identical ways (i.e., same factor loadings on measured items). Scalar invariance examined whether teachers with the same score of a latent construct (i.e., 6 behavioral domains) would be rated with the same observed scores (i.e., 32 high-inference items) when measured, irrespective of their country of origin (i.e., equal intercepts on measured item). Scalar invariance permits trustworthy comparisons of factor means across groups (Byrne, 2012).

Root-mean-square error of approximation (RMSEA), comparative fit index (CFI) and Tucker-Lewis index (TLI) are frequently used goodness of fit indices for categorical CFA and MGCFA models, and they all adhered to the same guidelines for a good model fit (i.e., CFI $\geq$ 0.90, TLI $\geq$ 0.90, SRMR $\leq$ 0.10, and RMSEA $\leq$ 0.08 are considered acceptable) (Hu and Bentler, 1999; Vandenberg and Lance, 2000). Additionally, the deterioration of model fit between successively constraint invariance models was examined by referring to the changes in CFI (ΔCFI), RMSEA (ΔRMSEA), and SRMR (ΔSRMR), with changes above 0.01

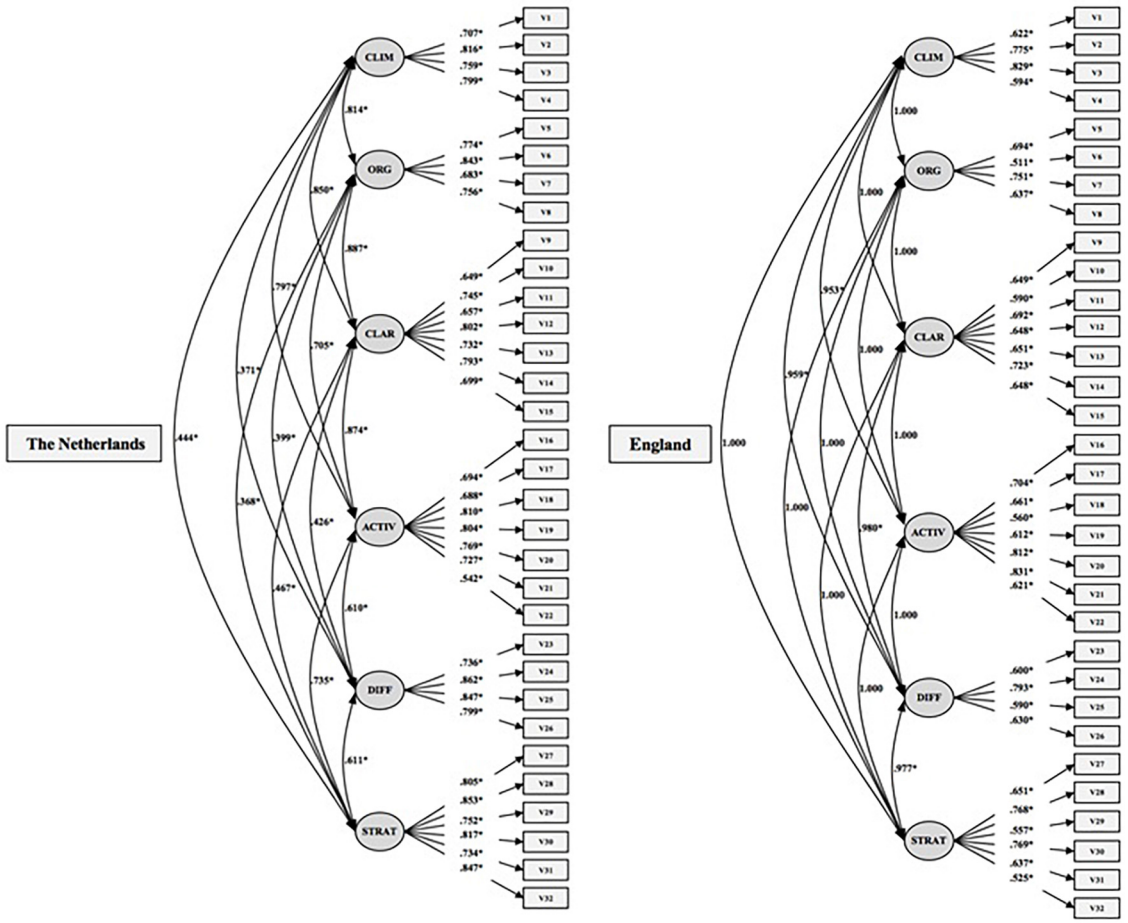


FIGURE 1
Factorial structures of effective teaching behavior for the Netherlands and England.

TABLE 5 Categorical multi-group confirmatory factor analysis for the Netherlands and England.

	Chi-square (df)	CFI	RMSEA with 90% CI	SRMR	Model comp	ΔCFI	ΔRMSEA	ΔSRMR	Decision
M1: Configural invariance	1918.894* (936)	0.887	0.082 [0.076, 0.087]	0.104					
M2: Metric invariance	1865.558*(968)	0.897	0.077 [0.071, 0.082]	0.111	M1	0.010	−0.005	0.007	Accepted
M3: Scalar invariance	2206.200* (994)	0.861	0.088 [0.083, 0.093]	0.115	M2	*−0.036*	*0.011*	0.004	Rejected
M3a: Partial scalar invariance[a]	1965.743* (986)	0.888	0.079 [0.074, 0.085]	0.111	M2	−0.009	0.002	0.000	Accepted
M4: Comparing latent means	1601.528* (942)	0.924	0.067 [0.061, 0.072]	0.097					

[a]Set free the thresholds of Item 2 in learning climate, Item 26 in differentiated Instruction, and Item 29 and 32 in teaching learning strategies. *$p < 0.05$.

indicating significant differences (Cheung and Rensvold, 2002; Chen, 2007).

5. Results

5.1. Preliminary analysis

Before investigating measurement invariance, the factor structure of effective teaching behavior in the three education contexts was investigated. If the hypothesized factor structure was confirmed using the empirical data, the investigation of measurement invariance was possible. The full measurement model of CFA displayed acceptable fit levels in the Netherlands and England, but not in the US (see Table 3). In the English model, 11 correlations between the latent constructs (i.e., 6 behavioral domains) were found above 1.00. After these correlation coefficients were adjusted to 1.00, the England model consistently showed acceptable fit indices. Comparing the adjusted model with the prior model revealed no significant changes in model fit. For the US data, the desired acceptable full measurement model could not be reached. One of the major issues was the cross-loadings of Item 1 ("The teacher shows respect for learners in his/her behavior and language"), Item 18 ("The teacher stimulates learners to think about solutions") and Item 4 ("The teacher fosters mutual respect"). In contrast to the positive domain correlations in the other two countries' models, the correlations between Learning Climate and Differentiated Instruction ($r = -0.355$, $p < 0.01$) and Teaching Learning Strategies ($r = -0.655$, $p < 0.001$) were moderately negative. Minor (e.g., error-term covariance) to major modifications (removal of 28% of the items) to the US data, which were made based on the modification indices, progressively improved the model-data fit, but not to the acceptable level. Therefore, the subsequent MGCFA only includes the full measurement models of the Netherlands and England.

All items sufficiently loaded on their corresponding domains, as shown by the factor loadings of all behavioral domains in the Netherlands and England being above the standard criterion of 0.40 (see Table 4 and Figure 1). Comparisons of the loadings across nations are not advised at this time since we have not yet established cross-contexts invariance.

5.2. Measurement invariance of teaching behavior

Although the six-factor structure was validated in each of the two education contexts separately, the results of categorical MGCFA using the two country data showed no convergence. A careful examination of the data detected instances of no celling filling in various item response categories of both countries (Netherlands: Item 1 with categories 1 and 2 unfilled, Item 2 and 3 with the category 1 unfilled; England: 20 items with the category 1 unfilled). Hence, we decided to re-categorize the value from 4 to 3 categories by collapsing the category 1 to 2.

Model fit indicators of MGCFA on the collapsed merged data are reported in Table 5. The configural model falls slightly under the minimum required standard. When the loadings were constraint to be identical (i.e., metric invariance), the change in fit was fairly minimal (ΔCFI = 0.010, ΔRMSEA = −0.005, and ΔSRMR = 0.007), showing an acceptable model-data fit relative to the configural model. For the scalar model, both CFI and RMSEA suggested a significant fit reduction (ΔCFI = −0.036, ΔRMSEA = 0.011, and ΔSRMR = 0.004). To obtain a better model fit (ΔCFI = −0.009, ΔRMSEA = 0.002, and ΔSRMR = 0.000), eight intercepts were released. This indicates that partial scalar invariance was supported for the two country's data. Partial scalar invariance is a sufficient criterion for comparing mean scores. No further modification or additional invariance models were included. A total of 28 items (87.5%) had invariant intercepts. The parameter estimates for this model are shown in Table 6.

5.3. Teaching behavior across the three education contexts

The comparison of latent means between the Netherlands and England was supported after achieving partial scalar invariance. Results revealed that teachers in England displayed lower levels of relatively basic behavioral domains (Stimulating Learning Climate and Classroom Management), but performed better on all relatively advanced domains (Activating Teaching, Differentiated Instruction, and Teaching Learning Strategies) ($p < 0.001$, see

TABLE 6 Standardized factor loadings for two countries in the partial scalar model.

Country and subscales	Standardized factor loadings							Domain correlations					Variance explained
								1	2	3	4	5	
Netherlands													
1. Stimulating teaching (four items)	0.687	0.802	0.779	0.728									56.3%
2. Classroom management (four items)	0.764	0.741	0.733	0.716				0.884					54.6%
3. Clarity of instruction (seven items)	0.661	0.703	0.688	0.758	0.740	0.771	0.695	0.905	0.958				51.5%
4. Activating teaching (seven items)	0.738	0.729	0.790	0.758	0.762	0.758	0.595	0.826	0.735	0.892			54.1%
5. Differentiated teaching (four items)	0.768	0.857	0.785	0.761				0.443	0.510	0.494	0.645		63.0%
6. Teaching learning strategies (six items)	0.797	0.857	0.715	0.806	0.740	0.748		0.563	0.447	0.551	0.815	0.596	60.6%
England													
1. Stimulating teaching (four items)	0.687	0.802	0.779	0.728									56.3%
2. Classroom management (four items)	0.764	0.741	0.733	0.716				0.890					54.6%
3. Clarity of instruction (seven items)	0.661	0.703	0.688	0.758	0.740	0.771	0.695	0.879	0.813				51.5%
4. Activating teaching (seven items)	0.738	0.729	0.790	0.758	0.762	0.758	0.595	0.887	0.865	0.925			54.1%
5. Differentiated teaching (four items)	0.768	0.857	0.785	0.761				0.740	0.697	0.762	0.831		63.0%
6. Teaching learning strategies (six items)	0.797	0.857	0.715	0.806	0.740	0.748		0.797	0.804	0.767	0.789	0.665	60.6%

All factor loadings and factor correlations reported in the table are significant (p <0.001).

10.3389/feduc.2023.1068938

Table 7 and Figure 2) compared to teachers in the Netherlands. There was no significant difference between the two education contexts reported for Clarity of Instruction. Among these behavioral domains, Differentiated Instruction showed the largest cross-context variation.

TABLE 7 Latent means of the partially scalar equivalent MGCFA model for the Netherlands and England.

	Netherlands (*N* =200)	England (*N* =115)
Learning climate	0.000	−0.603**
Classroom management	0.000	−0.742**
Clarity of instruction	0.000	−0.214
Activating teaching	0.000	0.759**
Differentiated instruction	0.000	2.559**
Teaching learning strategies	0.000	1.484**

**p <0.001. Netherlands is the reference category.

Examination of the mean scores revealed similar patterns with the MGCFA results, showing that teachers in the Netherlands showed higher levels of Learning Climate and Classroom Management, but lower levels of Activating Teaching, Differentiated Instruction and Teaching Learning Strategies compared to teachers in England (see Figure 3). Most notably, the largest difference was observed for Differentiated Instruction. Although it is tempting to say something about observed effective teaching behavior of teachers in the US, which based on mean scores seem to be generally lower compared that of teachers in the Netherlands and England, we have refrained from doing so due to the construct validity issue with the current US data and no evidence of measurement invariance.

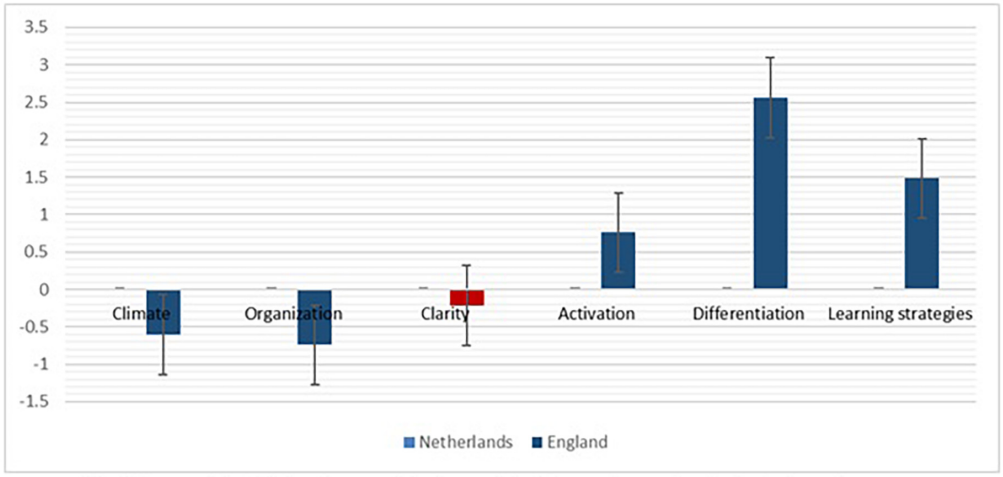


FIGURE 2
Latent means on the six domains between the Netherlands and England. The difference in clarity instruction in not significant. The Netherlands is reference category.

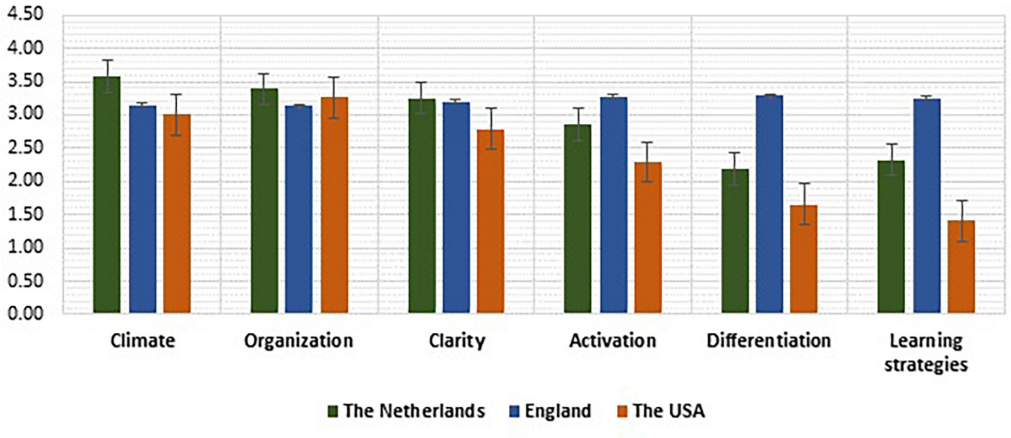


FIGURE 3
Raw scores means of the six domains. The six-factor structure is not supported in the USA data. Hence, the US result are not compared subsequently. Display is for general indication only.

6. Conclusion and discussion

The current study aimed to investigate measurement invariance of observer scoring of effective teaching behavior of secondary school teachers in three education contexts, including the Netherlands, England and the US, using a uniform ICALT observation instrument. It also aimed to identify differences in effective teaching behavior in the three education contexts. The work extends the previous work using a similar observation instrument (i.e., van de Grift et al., 2014; Maulana et al., 2021) by studying effective teaching behavior in the Netherlands, England, and the US.

The first guiding question was: *To what extent is there evidence of an invariant internal structure regarding effective teaching behavior in the Netherlands, England, and the United States?* We found that the six-factor structure of effective teaching behavior, as measured by the ICALT observation instrument, was confirmed in the Netherlands and England, but not in the US. Subsequently, the US data could not be included in the investigation of measurement invariance for comparing effective teaching behavior with the Netherlands and England. Prior to this study, face validity of the ICALT observation instrument was conducted including the experts from the three education contexts which confirmed the relevance of the six-factor structure of effective teaching behavior in their contexts. The current study provides further empirical evidence supporting the relevance of effective teaching behaviors and the usefulness of the ICALT instrument for measuring teaching behaviors in secondary education in the Netherlands and England.

Despite heavy modifications, the six-factor structure of teaching behavior was not confirmed in the current US data. Poor model-data fit indicates the data set did not match with the hypothesized model either because of the poor item quality or because of the poor understanding of the observers regarding the ICALT items. Although reasons for the poor model-data fit in the US data were unclear, further examinations of potential sources of bias may shed some light, including observation objects (real classroom vs. videoed classroom) and observer characteristics. In the Netherlands and England, teachers were observed and rated by trained observers in their classroom directly and in real-time, while in the US teachers were videotaped first and then rated by trained observers at a different time. The difference in the way teachers were observed (real-time actual classroom setting vs. videotapes) may explain differences in scoring teaching behavior. Furthermore, observers who observed and rated teachers in Netherlands and England were highly experienced (mostly senior) teachers or researchers, while observers who observed and rated teachers in the US, although highly knowledgeable in teaching behavior research, were relatively inexperienced teachers. In addition, the observers scoring teachers in the US are not from American backgrounds, while the observers scoring teachers in the Netherlands and England are from their original background. There is an indication that observers with different background interpret and answer the same items differently (Jansen et al., 2013). These differences may pose bias and affect the quality of the data (van de Vijver and Tanzer, 2004), and should be investigated in further research.

Measurement invariance of effective teaching behavior was established for the Netherlands and England, although full scalar invariance was not obtained. However, partial scalar invariance was a sufficient condition for comparing effective teaching behavior between groups (Steenkamp and Baumgartner, 1998). Valid inferences about the differences between latent factor means in the model can be made as long as there are at least two loadings and intercepts that are constrained equal across groups (Byrne et al., 1989). Based on the obtained partial scalar invariance, the latent factor means of effective teaching behavior in the Netherlands and England can be compared. However, comparing the sum scores or comparable observed means is not deemed acceptable because for this purpose full scalar invariance must be established (Steinmetz, 2013).

The second question was: *What are the differences regarding the six domains of effective teaching behavior across the three education contexts?* Due to the construct validity issue with the US data, and the fact that only partial scalar invariance was achieved for the Netherlands and England, this question can only be answered partially. Based on the latent factor means of scalar invariance, we found that teachers in England displayed lower levels of Learning Climate and Classroom Management behaviors compared to teachers in the Netherlands. On the contrary, English teachers showed higher levels of Activating Teaching, Differentiated Instruction and Teaching Learning Strategies compared to Dutch teachers. Learning Climate and Classroom Management are distinguished as relatively basic teaching skills, while Activating Teaching, Differentiated Instruction and Teaching Learning Strategies as relatively complex (van de Grift et al., 2014; Maulana et al., 2015). This finding is in line with findings of previous studies indicating that the majority of teachers in the Netherlands showed a good mastery of basic teaching skills (OECD, 2016; Inspectie van het Onderwijs, 2019). However, most teachers in the Netherlands still struggle in mastering complex teaching skills, particularly skills related to Differentiated Instruction (Maulana et al., 2020). Our finding seems to be in line with these past studies.

The 2018 TALIS report revealed that teachers in England and the Netherlands widely applied Classroom Management and Clarity of Instruction in their daily instructional practices, with the majority of teachers frequently dealing with disruptive students (OECD, 2019). In both education systems, complex teaching behaviors such as Teaching Learning Strategies were reported as less widespread compared to Classroom Management (OECD, 2019). Based on these TALIS results, there is evidence that teachers in both countries seem to commonly deal with basic teaching behavior in their daily practices, while exercising more complex behaviors seem to be less common. Our findings complement the TALIS findings by showing that although the two education systems share similar trends in teaching behavior practices as reported by teachers in the TALIS study, the current study shows that differences are evident between the two education contexts as reported by observers.

Most notably, the levels of Differentiated Instruction practices were observed to differ the most between the Netherlands and England. This can be related to differences in teacher preparation related to this teaching behavior between the two education contexts. In England, most teachers were trained to teach in mixed-ability settings as part of their formal initial teacher education (OECD, 2019). In the Netherlands, on the other hand, Differentiated Instruction is not widely included as an important

part of the initial teacher education curriculum yet, although much discussed in the current education agenda. Previous studies have showed that the quality of complex teaching behaviors, such as Differentiated Instruction and Teaching Learning Strategies were also observed to be low in other education contexts such as Republic of Korea, Indonesia, and South Africa (Maulana et al., 2021). Particularly, Differentiated Instruction was observed to be the lowest (Maulana et al., 2021). This indicates that Differentiated Instruction is not common practice in many countries, and remains a complex skill to master by many teachers (van der Lans et al., 2018).

Differences in effective teaching behavior between England and the Netherlands may also be related to teacher preparation characteristics. As described in the context of the study section, teachers in England must undertake formal training and be awarded Qualified Teacher Status (QTS). There is a strict performativity and accountability agenda in schools, which is overseen by the Ofsted inspectorate. In the Netherlands, there are two teaching qualifications: first degree (academic-focused) and second degree (practice-oriented) qualification. In addition, it is possible to teach in secondary education without teaching qualification as a second-career entrant or a guest teacher (Rijksoverheid, 2023), with the requirement to qualify within 2 years. These background differences between the two education contexts may explain why the Dutch sample shows higher levels of basic teaching behavior, while the English sample shows higher levels of more advance teaching behavior. Effective teaching behavior plays a central role in student learning and outcomes (e.g., Chapman et al., 2012; Hattie, 2012), and teachers' work matters the most for student outcomes (Hattie, 2012; van de Grift et al., 2017). Our findings seem to be in line with this general trend. Higher performances of English teachers in higher levels of teaching behaviors may correspond to higher performances of their students, particularly in reading and science, compared to Dutch teachers and their students (OECD, 2019).

7. Implications

Our study provides evidence that comparing effective teaching behavior using the ICALT observation instrument across different education systems is promising. However, establishing the factor structure of effective teaching behavior across all contexts may remain a first challenge in this endeavor. In our case, the factor structure was established quite well in the two education contexts (the Netherlands and England), but not in the third context (the US). We speculated that differences in rating objects (real classroom vs. video-taped classroom) and rater characteristics may play a role in explaining the failure to confirm the hypothesized factor structure in the US. This suggests that effective teaching behavior may be interpreted more similarly in certain contexts but not in other contexts, which implies that establishing the hypothesized factor structure in certain contexts will require more time and effort to potentially modify poor functioning items due to some cultural and practical differences.

Although the hypothesized factor structure is confirmed in the Netherlands and England data, reaching a full invariance of the measure in the two education contexts was a more difficult challenge. Nevertheless, reaching a partial scalar invariance for a rather complex measure of effective teaching behavior such as the ICALT is quite an accomplishment. Often, other measures of similar constructs hardly fulfill the requirement of scalar invariance (e.g., Muijs et al., 2018; OECD, 2019). This implies that the ICALT measures can be used to compare effective teaching behavior in the Netherlands and England, as long as the latent factor means are used for comparison (Steinmetz, 2013).

Based on similarities and differences in effective teaching behavior found between teachers in the Netherlands and England, implications for research and practices can be drawn. Mutual and reciprocal learning exchanges between the two education contexts are advocated. Teachers in the Netherlands can potentially learn from teachers in England regarding strategies and approaches for mastering higher levels of complex skills, particularly Activating Teaching, Differentiated Instruction and Teaching Learning Strategies. Teachers' Differentiated Instruction practices was particularly observed to differ the most between the two education contexts. This teaching skill has been particularly recommended in contemporary classroom practices, yet quite complex to master for most teachers in many countries (Maulana et al., 2023). Similarly, teachers in the English context can potentially learn from teachers in the Netherlands regarding strategies and approaches to improve basic teaching behavior skills related to Learning Climate and Classroom Management.

The TALIS 2018 study revealed that both teachers in the Netherlands and the UK reported a similar pattern and a comparable degree regarding a high prevalence of practicing more basic teaching behavior such as Classroom Management, but low widespread of practicing more complex behavior such as Teaching learning Strategies (OECD, 2019). In contrast, the current study revealed differences regarding basic and complex teaching behavior between the two education contexts. This implies that modes of collecting information (teacher report vs. observation) about teaching behavior may deliver divergent results, which confirm the necessity for doing triangulation (e.g., questionnaires, observation) to study effective teaching behavior. Observation offers value-added for unraveling differences in actual teaching behavior, which can be contrasted with questionnaire surveys for examining subjective perceptions of participants (Maulana and Helms-Lorenz, 2016).

8. Limitations and future directions

This study is subject to several limitations. First, although a random sampling method was initially planned, this was not realistic to employ. Subsequently, a more convenient sampling method was applied. Second, the sample size per country is relatively small. Given the costly and laborious nature of classroom observations, however, it was not feasible to observe more teachers. Third, observation was only done once for each measurement moment, which creates a snapshot of teaching behavior. Due to these limitations, generalizations of findings at country level is not recommended until more representative and more randomly selected samples are available.

Fourth, although the observer training in the three education contexts was applied using identical procedures and standards, and we made sure that only observers who passed the extensive

training were invited to observe classrooms, we have no control over the observation quality of the observers in practice. Cultural influences and day-to-day context-specificity may affect the quality of observation. Future studies should develop a control mechanism during the actual observation to minimize bias and improve data quality, if possible. It may also be worthwhile to employ cross-observer designs across education contexts to minimize cultural bias in observation (Maulana et al., 2021).

Finally, it was not possible to observe teachers in their natural settings directly in the US due to some resource problems. Hence, available video-taped lessons were used as an alternative strategy. It may be possible that observers rate teachers differently when observing actual lessons compared to video-taped lessons using the ICALT instrument. This speculation should be investigated further to establish whether raters and/or the ICALT instrument is sensitive to differences regarding the object of observation. In addition, it is coincident that the US data was rated by less experienced teachers. Failing to confirm the hypothesized factor structure of effective teaching behavior in this education context may partially be explained by this observer characteristics, which should be validated in future research.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Commission of Teacher Education University of Groningen, Worcester, and the Education University of Hong Kong. The patients/participants provided their written informed consent to participate in this study.

Author contributions

RM conceived and designed the study, wrote the manuscript, checked statistical analyses, and coordinated the manuscript. MH-L co-designed the study and contributed to the manuscript. AK, JK, BL, KH-M, and KB wrote sections of the manuscript. XF performed statistical analyses and wrote sections of the manuscript. All authors read and approved the submission of the manuscript.

Funding

This work was supported by the Dutch scientific funding agency (NRO) under Grant number 405-15-732.

Acknowledgments

We are indebted to Peter Moorer for his assistance with the data management, and to Jaap Buitink, Geke Schuurman, and Carla Griep for their valuable contributions as international trainers. The contribution of Steve Cantrell in assisting with the MET data access is highly acknowledged. We also thank Wang Xiuping for her contribution with the manuscript. The biggest thanks goes to all teachers, schools, and observers participating in this large-scale study in the three countries.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor AL declared a past collaboration with the author AK.

Publisher's note

References

Adamson, B. (2012). International comparative studies in teaching and teacher education. *Teach. Teach. Educ.* 28, 641–648. doi: 10.1016/j.tate.2012.02.003

Bae, H., and Kwon, K. (2021). Developing metacognitive skills through class activities: what makes students use metacognitive skills? *Educ. Stud.* 47, 456–471. doi: 10.1080/03055698.2019.1707068

Bell, C. A., Dobbelaer, M. J., Klette, K., and Visscher, A. (2019). Qualities of classroom observation systems. *Sch. Eff. Sch. Improv.* 30, 3–29. doi: 10.1080/09243453.2018.1539014

Bolkan, S. (2016). The Importance of Instructor Clarity and Its Effect on Student Learning: Facilitating Elaboration by Reducing Cognitive Load. *Commun. Rep.* 29, 152–162. doi: 10.1080/08934215.2015.1067708

Bray, M., Adamson, B., and Mason, M. (2014). "Different models, different emphases, different insights," in *Comparative education research*, eds M. Bray, B. Adamson, and M. Mason (Cham: Springer), 417–436. doi: 10.1007/978-3-319-05594-7_15

Brouwer, J., Jansen, E., Flache, A., and Hofman, A. (2016). The impact of social capital on self-efficacy and study success among first-year university students. *Learn. Individ. Differ.* 52, 109–118. doi: 10.1016/j.lindif.2016.09.016

Bruns, B., and Luque, J. (2015). *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. Washington, DC: World Bank. doi: 10.1596/978-1-4648-0151-8

Byrne, B. M. (2012). *Structural equation modeling with Mplus: basic concepts, applications, and programming*. New York, NY: Routledge. doi: 10.4324/9780203807644

Byrne, B. M., Shavelson, R. J., and Muthen, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychol. Bull.* 105, 456–466. doi: 10.1037/0033-2909.105.3.456

Chapman, C., Harris, A., Armstrong, P., and Chapman, C. (2012). *School effectiveness and improvement research, policy and practice*. London: Taylor & Francis. doi: 10.4324/9780203136553

Chapman, J., Schetzsle, S., and Wahlers, R. (2016). An innovative, experiential-learning project for sales management and professional selling students. *Mark. Educ. Rev.* 26, 45–50. doi: 10.1080/10528008.2015.1091674

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equat/ Model.* 14, 464–504. doi: 10.1080/10705510701301834

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *Am. Econ. Rev.* 104, 2593–2632. doi: 10.1257/aer.104.9.2593

Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equat. Model.* 9, 233–255. doi: 10.1207/S15328007SEM0902_5

Coe, R., Aloisi, C., Higgins, S., and Major, L. (2014). *What makes great teaching? Review of the underpinning research*. London: Sutton Trust.

Commission of Teacher Credentialing of California (2021). *Preparation Programs*. Sacramento, CA: Commission of Teacher Credentialing of California.

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Rev. Educ. Res.* 77, 113–143. doi: 10.3102/00346543029 8563

Creemers, B. P. M., and Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice, and theory in contemporary schools*. London: Routledge. doi: 10.4324/9780203939185

Danielson, C. (2007). *Enhancing professional practice: A framework for teaching*, 2nd Edn. Alexandria, VA: Association for Supervision & Curriculum Development.

Danielson, C. (2013). *The framework for teaching evaluation instrument*. Chicago, IL: The Danielson Group.

Dulfer, N., and Akhlaghi Koopaei, F. (2021). Moving from Feedback to Feedforward in IBDP classrooms. *J. Res. Int. Educ.* 20, 101–113. doi: 10.1177/14752409211032528

Fuller, F. (1970). *Personalized education for teachers: One application of the teachers concerns model*. Austin, TX: R&D Center for Teacher Education.

Ginner Hau, H., Ferrer-Wreder, L., and Allodi, M. W. (2021). "Capitalising on Classroom Climate to Promote Positive Development," in *Handbook of Positive Youth Development*. *Springer Series on Child and Family Studies*, eds R. Dimitrova and N. Wiium (Cham: Springer). doi: 10.1007/978-3-030-70262-5_25

Goodwin, A. L., and Low, E. L. (2021). Rethinking conceptualisations of teacher quality in Singapore and Hong Kong: A comparative analysis. *Eur. J. Teach. Educ.* 44, 365–382. doi: 10.1080/02619768.2021.1913117

Gov.uk (2021). *Schools, pupils and their characteristics*. London: National Statistics.

Grossman, P., Loeb, S., Cohen, J. J., and Wyckoff, J. (2013). Measure for measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. *Am. J. Educ.* 119, 445–470. doi: 10.1086/669901

Hanushek, E. A., and Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *Am. Econ. Rev.* 100, 267–271. doi: 10.1257/aer.100.2.267

Harris, D. N., and Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *J. Public Econ.* 95, 798–812. doi: 10.1016/j.jpubeco.2010.11.009

Hattie, J. (2012). *Visible learning for teachers: Maximising impact on learning*. London: Routledge. doi: 10.4324/9780203181522

Hattie, J., and Clinton, J. (2008). "Identifying accomplished teachers: A validation study," in *Advances in program evaluation: Vol. 11. Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards*, eds L. Ingvarson and J. Hattie (Bingley: Emerald Group Publishing), 313–344. doi: 10.1016/S1474-7863(07)11011-5

Helms-Lorenz, M., van der Pers, M., Moorer, P., Harmsen, R., and Verkade, A. (2019). *Inductie in het Noorden: Eindrapportage 2014-2019*. Groningen: University of Groningen.

Helms-Lorenz, M., van der Pers, M., Moorer, P., Lugthart, E., van der Lans, R., and Maulana, R. (2020). *Begeleiding Startende Leraren 2014-2019*. Noorden: Eindrapportage.

Hennessy, S., Howe, C., Mercer, N., and Vrikki, M. (2020). Coding classroom dialogue: Methodological considerations for researchers. *Learn. Cult. Soc. Interact.* 25, 100404. doi: 10.1016/j.lcsi.2020.100404

Houtveen, A. A. M., and van de Grift, W. J. C. M. (2007). Effects of metacognitive strategy instruction and instruction time on reading comprehension. *Sch. Eff. Sch. Improv.* 18, 173–190. doi: 10.1080/09243450601058717

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equat. Model.* 6, 1–55. doi: 10.1080/10705519909540118

Inspectie van het Onderwijs (2018). *The state of education in the Netherlands 2016/2017*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs (2019). *The state of education in the Netherlands 2018*. Utrecht: Inspectie van het Onderwijs.

Irnidayanti, Y., Maulana, R., Helms-Lorenz, M., and Fadhilah, N. (2020). Relationship between teaching motivation and teaching behaviour of secondary education teachers in Indonesia. *J. Study Educ. Dev.* 43, 271–308. doi: 10.1080/02103702.2020.1722413

Ismajli, H., and Imami-Morina, I. (2018). 'Differentiated Instruction: Understanding and Applying Interactive Strategies to Meet the Needs of all the Students. *Int. J. Instruct.* 11, 207–218. doi: 10.12973/iji.2018.11315a

Jansen, E., André, S., and Suhre, C. (2013). Readiness and expectations questionnaire: A cross-cultural measurement instrument for first-year university students. *Educ. Assess. Eval. Account.* 25, 115–130. doi: 10.1007/s11092-013-9161-2

Jortveit, M., Tveit, A. D., Cameron, D. L., and Lindqvist, G. (2020). A comparative study of Norwegian and Swedish special educators' beliefs and practices. *Eur. J. Spec. Needs Educ.* 35, 350–365. doi: 10.1080/08856257.2019.1689716

Kindsvatter, R., Wilen, W., and Ishler, M. (1988). *Dynamics of effective teaching*. Harlow: Longman.

Kington, A., Sammons, P., Brown, E., Regan, E., Ko, J., and Buckler, S. (2014). *EBOOK: Effective Classroom Practice*. London: McGraw-Hill Education.

Klette, K. (2022). "The use of Video Capturing in International Large-Scale Assessment Studies: Methodological and Theoretical Considerations," in *International Handbook of Comparative Large Scale Studies in Education: Perspectives, Methods and Findings*, eds T. Nilsen, A. Stancel-Piatak, and J.-E. Gustafsson (Cham: Springer International Publishing), 470–510. doi: 10.1007/978-3-030-88178-8_19

Klieme, E., Pauli, C., and Reusser, K. (2009). "The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms," in *The power of video studies in investigating teaching and learning in the classroom*, eds T. Janík and T. Seidel (Halifax: Waxmann), 137–160.

Ko, J., and Sammons, P. (2013). *Effective teaching: A review of research and evidence*. London: CfBT Education Trust.

Kraft, M. A., and Lyon, M. A. (2022). *The Rise and Fall of the Teaching Profession: Prestige, Interest, Preparation, and Satisfaction over the Last Half Century. (Ed Working Paper: 22-679)*. Providence, RI: Annenberg Institute at Brown University.

Luoto, J. M. (2023). Comparative education and comparative classroom observation systems. *Comp. Educ.* [Epub ahead of print]. doi: 10.1080/03050068.2023.2173917

Martens, M., and Asbrand, B. (2022). "Documentary Classroom Research. Theory and Methodology," in *Dokumentarische Unterrichtsforschung in den Fachdidaktiken: Theoretische Grundlagen und Forschungspraxis*, eds M. Martens, B. Asbrand, T. Buchborn, and J. Menthe (Berlin: Springer), 19–38. doi: 10.1007/978-3-658-32566-4_2

Martin, M. O., Mullis, I. V., and Hooper, M. (2016c). *Methods and procedures in TIMSS 2015*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Martin, M. O., Mullis, I. V. S., Foy, P., and Hooper, M. (2016a). *TIMSS 2015 international results in science*. Boston, MA: TIMSS.

Martin, M. O., Mullis, I. V. S., Foy, P., and Hooper, M. (2016b). *TIMSS achievement ethodology. In Methods and procedures in TIMSS 2015*. Boston, MA: TIMSS.

Marzano, R. J. (2003). *What works in schools. Translating research into action*. Arlington, VA: ASCD.

Maulana, R., André, S., Helms-Lorenz, M., Ko, J., Chun, S., Shahzad, A., et al. (2021). Observed teaching behaviour in secondary education across six countries: measurement invariance and indication of cross-national variations. *Sch. Effective. Sch. Improve.* 32, 64–95. doi: 10.1080/09243453.2020.1777170

Maulana, R., Feng, X., Helms-Lorenz, M., Ko, J., Chun, S., Abid, S., et al. (2022). *Observing Teaching Behavior Using the International Comparative Analysis of Learning and Teaching Measure Across Countries: Is There Measurement Invariance?*. San Diego, CA: University of Worcester.

Maulana, R., and Helms-Lorenz, M. (2016). Observations and student perceptions of the quality of preservice teachers' teaching behaviour: Construct representation and predictive quality. *Learn. Environ. Res.* 19, 335–357. doi: 10.1007/s10984-016-9215-8

Maulana, R., Helms-Lorenz, M., Moorer, P., Smale-Jacobse, A. E., and Feng, X. (2023). *Differentiated Instruction in Teaching from the International Perspective: Methodological and practical insights*. Groningen: University of Groningen Press, doi: 10.21827/62c5541759973

Maulana, R., Helms-Lorenz, M., and van de Grift, W. (2015). A longitudinal study of induction on the acceleration of growth in teaching quality of beginning teachers through the eyes of their students. *Teach. Teach. Educ.* 51, 225–245. doi: 10.1016/j.tate.2015.07.003

Maulana, R., Helms-Lorenz, M., and van de Grift, W. (2017). Validating a model of effective teaching behaviour of pre-service teachers. *Teach. Teach.* 23, 471–493. doi: 10.1080/13540602.2016.1211102

Maulana, R., Opdenakker, M.-C., and Bosker, R. (2014). Teacher–student interpersonal relationships do change and affect academic motivation: A multilevel growth curve modelling. *Br. J. Educ. Psychol.* 84, 459–482. doi: 10.1111/bjep.12031

Maulana, R., Smale-Jacobse, A., Helms-Lorenz, M., Chun, S., and Lee, O. (2020). Measuring differentiated instruction in the Netherlands and South Korea: Factor structure equivalence, correlates, and complexity level. *Eur. J. Psychol. Educ.* 35, 881–909. doi: 10.1007/s10212-019-00446-4

Millsap, R. E., and Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multiv. Behav. Res.* 39, 479–515. doi: 10.1207/S15327906MBR3903_4

Moberg, S., Muta, E., Korenaga, K., Kuorelahti, M., and Savolainen, H. (2020). Struggling for inclusive education in Japan and Finland: teachers' attitudes towards inclusive education. *Eur. J. Spec. Needs Educ.* 35, 100–114. doi: 10.1080/08856257.2019.1615800

Mortimore, P., Sammons, P., Stoll, L., Lewis, D., and Ecob, R. (1988). *School matters*. Sacramento, CA: University of California Press.

Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H., and Earl, L. (2014). State of the art - teacher effectiveness and professional learning. *Sch. Eff. Sch. Improv.* 25, 231–256. doi: 10.1080/09243453.2014.885451

Muijs, D., Reynolds, D., Sammons, P., Kyriakides, L., Creemers, B. P. M., and Teddlie, C. (2018). Assessing individual lessons using a generic teacher observation instrument: How useful is the International System For Teacher Observation And Feedback (ISTOF)? *ZDM Mathe. Educ.* 50, 395–406. doi: 10.1007/s11858-018-0921-9

Mullis, I. V., Martin, M. O., Foy, P., and Hooper, M. (2016). *TIMSS 2015 international results in science*. Chestnut Hill: TIMSS & PIRLS International Study Centre.

Muthén, B., and Muthén, L. (2017). *Mplus. In Handbook of item response theory*. Boca Raton, FL: Chapman and Hall/CRC, 507–518.

Noben, I., Maulana, R., Deinum, J. F., and Hofman, W. H. (2021). Measuring university teachers' teaching quality: a Rasch modelling approach. *Learn. Environ. Res.* 24, 87–107. doi: 10.1007/s10984-020-09319-w

OECD (2013). *Country note: Results from TALIS 2013: United States of America*. Paris: OECD.

OECD (2014). *Education at a Glance 2014: OECD Indicators*. Paris: OECD.

OECD (2016). *Education at a glance 2016*. Paris: OECD.

OECD (2018). *Education at a Glance 2018: OECD Indicators*. Paris: OECD. doi: 10.1787/eag-2018-en

OECD (2019). *PISA 2018 Results: Combined Executive Summaries Volume I, II, III*. Paris: OECD.

OECD (2020). *Global Teaching InSights: A Video Study of Teaching*. Paris: OECD. doi: 10.1787/20d6f36b-en

OECD (2021b). *Global teaching insights: Technical report*. Paris: OECD.

OECD (2021a). *21st-Century Readers: Developing Literacy Skills in a Digital World*. Paris: OECD.

Ofsted (2021). *Education Inspection Framework*. London: Ofsted.

Opdenakker, M.-C., and Minnaert, A. (2011). Relationship between learning environment characteristics and academic engagement. *Psychol. Rep.* 109, 259–284. doi: 10.2466/09.10.11.PR0.109.4.259-284

Paas, F., and van Merriënboer, J. J. G. (2020). Cognitive-Load Theory: Methods to Manage Working Memory Load in the Learning of Complex Tasks. *Curr. Dir. Psychol. Sci.* 29, 394–398. doi: 10.1177/0963721420922183

Pianta, R. C., and Hamre, B. K. (2009). Conceptualisation, measurement, and improvement of classroom processes: Standardised observation can leverage capacity. *Educ. Res.* 38, 109–119. doi: 10.3102/0013189X09332374

Pianta, R. C., Hamre, B. K., and Mintz, S. L. (2010). *Classroom Assessment Scoring System (CLASS): Upper elementary manual*. Charlottesville, VA: Teachstone.

Powell, J. J. (2020). Comparative education in an age of competition and collaboration. *Comp. Educ.* 56, 57–78. doi: 10.1080/03050068.2019.1701248

Reis, S. M., McCoach, D. B., Little, C. A., Muller, L. M., and Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *Am. Educ. Res. J.* 48, 462–501. doi: 10.3102/0002831210382891

Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., and Schaffer, G. (2002). *World class schools: International perspectives in school effectiveness*. London: Routledge Falmer. doi: 10.4324/9780203164624

Rijksoverheid. (2023). *Hoe word ik leraar in het voortgezet onderwijs [How can I be a teacher in secondary education?]*. Den Haag: Rijksoverheid.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *Am. Econ. Rev.* 94, 247–252. doi: 10.1257/0002828041302244

Rosenshine, B. (2010). Principles of instruction; Educational practices series. *Int. Acad. Educ.* 21, 376–391.

Rosenshine, B. (2012). Principles of Instruction: Research-Based Strategies That All Teachers Should Know. *Am. Educ.* 36, 12–39.

Rosenshine, B., and Stevens, R. (1986). "Teaching functions," in *Handbook of research on teaching*, 3rd Edn, ed. M. C. Wittrock (Berlin: Macmillan), 376–391.

Scheerens, J., and Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon Press.

Sensevy, G. (2014). Characterizing teaching effectiveness in the Joint Action Theory in Didactics: An exploratory study in primary school. *J. Curricul. Stud.* 46, 577–610. doi: 10.1080/00220272.2014.931466

Simonsen, B., Freeman, J., Myers, D., Dooley, K., Maddock, E., Kern, L., et al. (2020). The Effects of Targeted Professional Development on Teachers' Use of Empirically Supported Classroom Management Practices. *J. Pos. Behav. Interv.* 22, 3–14. doi: 10.1177/1098300719859615

Smith, T., Baker, W., Hattie, J., and Bond, L. (2008). "A validity study of the certification system of the national board for professional teaching standards," in *Assessing teachers for professional certification: The first decade of the national board for professional teaching standards*, Vol. 11, eds L. C. Ingvarson, and J. Hattie (Amsterdam: Elsevier Press), 345–380.

Steenkamp, J. B. E., and Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *J. Consum. Res.* 25, 78–90. doi: 10.1086/209528

Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology* 9, 1–12. doi: 10.1027/1614-2241/a000049

Stroet, K., Opdenakker, M. C., and Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educ. Res. Rev.* 9, 65–87. doi: 10.1016/j.edurev.2012.11.003

Suter, L. E. (2019). "The status of comparative education research in the 21st century: An Empiricist's views," in *The SAGE Handbook of Comparative Studies in Education*, eds L. E. Suter, E. Smith, and B. D. Denman (London: SAGE), 3–24. doi: 10.4135/9781526470379.n2

Teddlie, C., Virgilio, I., and Oescher, J. (1990). Development and validation of the Virgilio Teacher Behavior instrument. *Educ. Psychol. Meas.* 50, 421–430.

Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., et al. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of literature. *J. Educ. Gifted* 27, 119–145. doi: 10.1177/016235320302700203

U.S. Department of Education (2013). *The condition of education* 2013. Available online at: https://nces.ed.gov/pubs2013/2013037.pdf

van de Grift, W., Helms-Lorenz, M., and Maulana, R. (2014). Teaching skills of student teachers: Calibration of an evaluation instrument and its value in predicting student academic engagement. *Stud. Educ. Eval.* 43, 150–159. doi: 10.1016/j.stueduc.2014.09.003

van de Grift, W. J. (2007). Quality of teaching in four European countries: A review of the literature and application of an assessment instrument. *Educ. Res.* 49, 127–152. doi: 10.1080/00131880701369651

van de Grift, W. J. (2014). Measuring teaching quality in several European countries. *Sch. Effective. Sch. Improve.* 25, 295–311. doi: 10.1080/09243453.2013.794845

van de Grift, W. J., Chun, S., Maulana, R., Lee, O., and Helms-Lorenz, M. (2017). Measuring teaching quality and student engagement in South Korea and The Netherlands. *Sch. Effective. Sch. Improve.* 28, 337–349. doi: 10.1080/09243453.2016.1263215

van de Vijver, F., and Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Eur. Rev. Appl. Psychol.* 54, 119–135. doi: 10.1016/j.erap.2003.12.004

van der Lans, R., Maulana, R., Helms-Lorenz, M., Fernández-García, C., Chun, S., de Jager, T., et al. (2021). Student perceptions of teaching quality in five countries: A Partial Credit Model approach to assess measurement invariance. *SAGE Open* 11, 21582440211040121. doi: 10.1177/21582440211040121

van der Lans, R. M., van de Grift, W. J. C. M., and van Veen, K. (2018). Developing an instrument for teacher feedback: Using the Rasch model to explore teachers' development of effective teaching strategies and behaviors. *J. Exp. Educ.* 86, 247–264. doi: 10.1080/00220973.2016.1268086

Vandenberg, R. J., and Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* 3, 4–70. doi: 10.1177/109442810031002

Wang, M. C., Reynolds, M. C., and Walberg, H. J. (1995). *Handbook of special and remedial education: Research and practice*, 2nd Edn. Oxford: Pergamon Press.

World Bank (2022). *Teach secondary: Helping countries track and improve teaching quality*. Washington, DC: World Bank.

Xing, C., and Hall, J. A. (2015). Confirmatory factor analysis and measurement invariance testing with ordinal data: An application in revising the flirting styles inventory. *Commun. Methods Meas.* 9, 123–151. doi: 10.1080/19312458.2015.1061651

Yair, G. (2000). Not just about time: Instructional practices and productive time in school. *Educ. Adm. Quart.* 36, 485–512. doi: 10.1177/00131610021969083

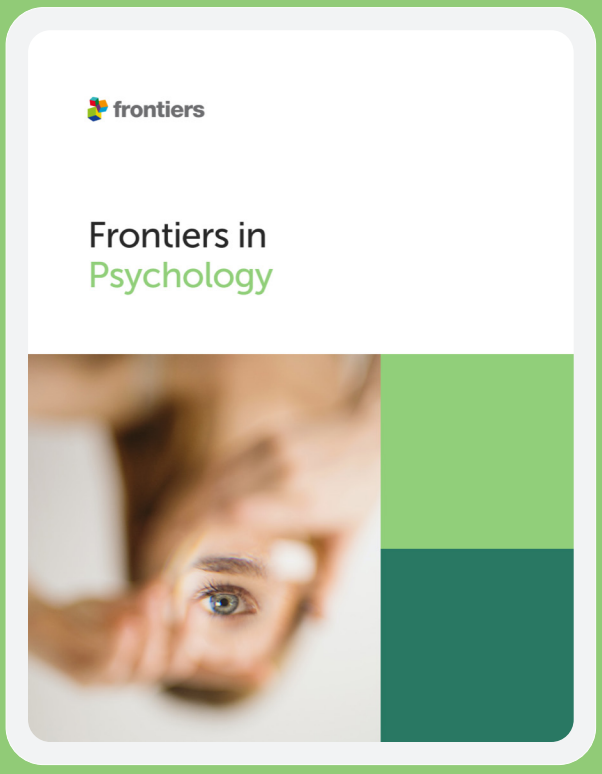
Frontiers in Psychology

Paving the way for a greater understanding of human behavior

The most cited journal in its field, exploring psychological sciences - from clinical research to cognitive science, from imaging studies to human factors, and from animal cognition to social psychology.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact