Frontline Learning Research
An official journal of EARLI

Frontline Learning Research Vol. 12 No. 1 (2024) 66 - 123
ISSN 2295-3159

Revisiting the Three Basic Dimensions model: A critical empirical investigation of the indirect effects of student-perceived teaching quality on student outcomes

Ayşenur Alp Christ[1*], Vanda Capon-Sieber[1*],

Carmen Köhler[2], Eckhard Klieme[2] & Anna-Katharina Praetorius[1]

[1] University of Zurich, Switzerland
[2] Leibniz Institute for Research and Information in Education (DIPF), Germany

*The first two authors have a shared contribution in conceptualization, writing, and revising the manuscript and should be considered co-first authors.

Article received 20 August 2023 / Article revised 30 November 2023 / Accepted 8 February 2024 / Available online 8 March

Abstract

The Three Basic Dimensions model, theorizes three mediators for the effect of teaching quality dimensions on student outcomes. However, the proposed mediating paths and their effects have largely not been empirically tested. This study investigated the mediating role of depth-of-processing, time-on-task, and need satisfaction between student-perceived teaching quality and student mathematics achievement and interest, expanding the TBD model to include mediation paths suggested by theories of motivation, cognition, and effort. Data from the TALIS Video Study for Germany, comprising 958 secondary school students in 41 classrooms, were used to run multilevel longitudinal and correlational mediation analyses. The results only found mediation effects at the student level; there were no mediating effects at the classroom level. Not all of the hypothesized relationships thought to exist between the mediators and achievement and interest outcomes were confirmed. The conceptual sequence of the variables, the choice of correlational vs. longitudinal evidence, and the level of analysis were all shown to have an impact on the results. The study thus confirms some of the assumptions of the TBD model, identifies new paths between teaching quality and student outcomes, and provides suggestions for how to proceed with further investigation of a model which should be expanded and more thoroughly empirically tested.

Keywords: Teaching quality, learning processes, mediation, interest, achievement, three basic dimensions

Corresponding authors: Ayşenur Alp Christ, Freiestrasse 36, 8032 Zürich, Switzerland, aysenur.alpchrist@uzh.ch & Vanda Capon-Sieber, Freiestrasse 36, 8032 Zürich, Switzerland, vanda.sieber@ife.uzh.ch DOI: *https://doi.org/10.14786/flr.v12i1.1349*

1. Introduction

The Three Basic Dimensions (TBD) model of teaching quality is influential and widely used by researchers in the field of teaching quality, particularly in German-speaking countries (Klieme et al., 2006, 2009; Kunter & Trautwein, 2013; Praetorius et al., 2018; Reusser et al., 2010). This model, developed by Klieme et al. (2001, 2009), offers a concise framework for understanding the aspects of teaching quality by categorizing them into three key dimensions: cognitive activation, classroom management, and student support. One of its principal advantages over the many other models and frameworks is that it integrates student learning processes, focusing on the mediating role that they play between teaching quality and student outcomes. For example, it hypothesizes that depth of processing mediates between cognitive activation and student achievement, which suggests that cognitive activation only has a significant impact on learning outcomes when students engage in deep processing (Klieme et al., 2006, see Figure 1). However, researchers have rarely conducted systematic empirical examinations of the assumed mediators. Although the role of mediators has been supported by results from studies focusing on specific paths in the model (for cognitive activation, e.g., Hiebert & Grouws, 2007; Stein & Lane, 1996; for classroom management, e.g., Hospel & Galand, 2016; Kunter, et al., 2007; and for student support, e.g., Kiemer et al., 2018; Mouratidis et al., 2013), overall, empirical research on these mediators remains very limited.

The incorporation of mediators between teaching quality and student outcomes in the TBD model was guided by selected theoretical considerations, primarily rooted in Self-Determination Theory (SDT, Ryan & Deci, 2000) and constructivism (De Corte, 2004; Pauli & Reusser, 2006). This selection of references may miss other valid theoretical perspectives that might explain the mediators and observed student outcomes. Moreover, the reasoning behind the choice of theory to explain mediation pathways in the TBD model is not well-articulated in the literature. The lack of clarity creates potential gaps in our understanding of the model and could lead to an incomplete representation of the role of mediators between teaching quality and student outcomes. It is therefore important to consider the possibility of additional theoretically likely relations between the variables in the model. For example, in the context of the original model, depth of processing is influenced by cognitive activation and classroom management. However, by incorporating theoretical insights from other theories, such as the elaboration likelihood model (ELM, Petty & Cacioppo, 1986), we propose that elements of student support, such as activities that accentuate the relevance of tasks, may also contribute to increased depth of processing.

This paper aims to address these gaps in the theoretical and empirical foundations of the TBD model. It seeks to comprehensively test the assumptions of the model, including additional possible mediation pathways, to provide a more robust understanding of the relationship between teaching quality and student outcomes.

1.1 The Three Basic Dimensions model of teaching quality

The TBD model (Figure 1) identifies cognitive activation, classroom management, and student support as the key aspects of teaching quality that affect student outcomes such as achievement and motivation. In particular, cognitive activation and classroom management are assumed to have an effect on student achievement and student support is linked to student motivation. The results of an empirical analysis conducted by Klieme et al. (2001) provide support for this idea. Emphasizing the role of student understanding, attentiveness, and motivation in the learning process (Diederich & Tenorth, 1997), the basic dimensions have been theoretically linked to students' *depth of processing, time-on-task,* and *need satisfaction* (i.e., student use of learning opportunities) (Klieme et al., 2006, 2009; Klieme & Rakoczy, 2008; see Figure 1). Specifically, it has been hypothesized that cognitive activation is linked to depth of processing, that student support has an effect on need satisfaction, and that classroom management is linked to depth of processing, time-on-task, and need satisfaction. For simplicity and parsimony, the original TBD model focused on pathways that included mediators between teaching quality and student outcomes based on the specific theoretical considerations used to formulate the basic dimensions (De

Corte, 2004; Pauli & Reusser, 2006; Ryan & Deci, 2000). To better explain how these dimensions relate to the use of learning opportunities by students, we first describe the basic dimensions in Section 1.1.2, then explain the mediators and outcomes in detail.

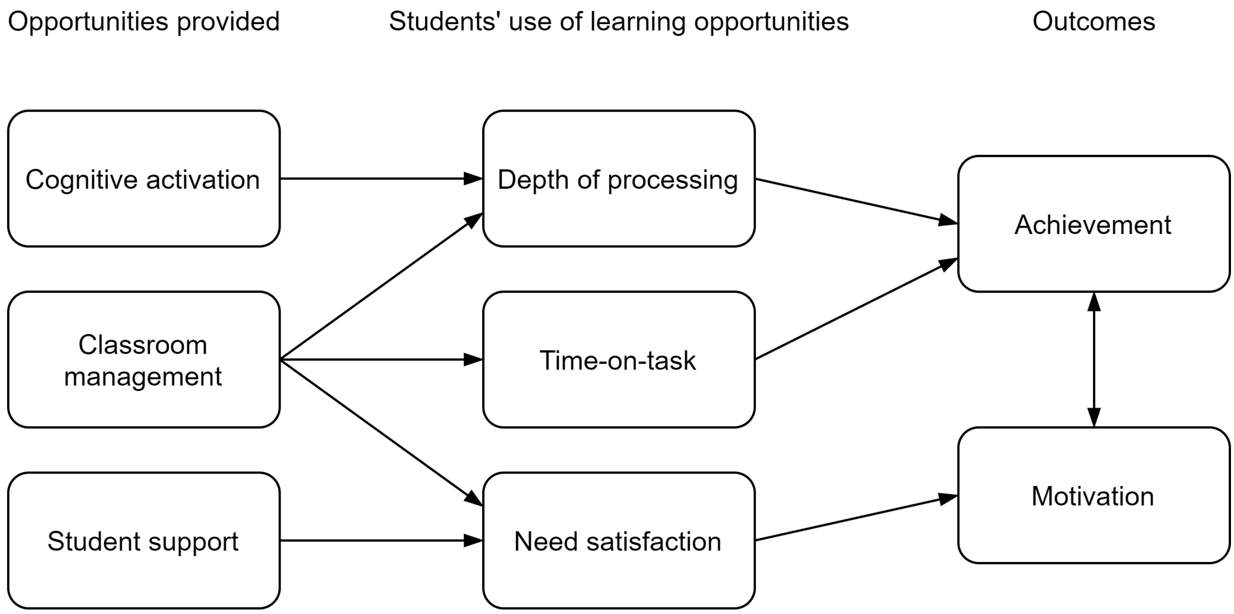

Figure 1. The relations between the three basic dimensions and student achievement and motivation according to the TBD model (adapted from Klieme et al., 2009).

1.1.2 The three basic dimensions of teaching quality

Cognitive activation. This dimension is based on different (socio-)constructivist learning theories (Aebli, 2011; Piaget, 1992; Vygotsky, 1978), which emphasize the independent construction of knowledge and interaction with others within the zone of proximal development (ZPD, Vygotsky, 1978). The current understanding of cognitive activation encompasses multiple facets that aim to stimulate higher-order cognitive processes (Lipowsky & Hess, 2019; Ziegelbauer, 2009). This includes encouraging students to understand learning content by providing challenging tasks and to activate prior knowledge, practicing content-related discourse, and fostering active participation in critical class discussions (Förtsch et al., 2018; Klieme et al., 2009; Lipowsky et al., 2009; Lotz, 2016; Praetorius et al., 2014, 2018; Rakoczy & Pauli, 2006). Some authors propose that the teaching behaviors should support students' independent engagement with the learning content (Lotz, 2016), aspects of self-regulation and metacognition (Praetorius et al., 2018; Rieser et al., 2016).

Classroom management. This dimension encapsulates the "effective strategies for organizing classrooms" proposed by several researchers (Doyle, 1986; Emmer & Stough, 2001; Evertson, 1989; Kounin, 1970a, 1970b; Kunter et al., 2007). The strategies result in increased learning time. This is, among others, the result of the "withitness" of a teacher, which means that a teacher is omnipresent during a lesson and informed about all that is happening in the classroom. With efficient time use, making effective transitions between topics and having clear rules and routines, a teacher can ensure the smooth running of the classroom. Successful classroom management also includes early, prompt, intervention to prevent disruptions and discipline problems (Kounin, 1970a; Kuger, 2016; Praetorius et al., 2018).

Student support. This dimension is based on SDT (Ryan & Deci, 2000, 2017) and comprises the support of student competence, autonomy, and relatedness (Klieme et al., 2009). Student support includes giving constructive feedback, addressing student errors and misconceptions in a positive manner, and nurturing an atmosphere of mutual care and respect in the classroom (see Fauth et al., 2014, 2019; Lipowsky et al., 2009; Praetorius et al., 2018). According to SDT, it also involves understanding student needs, helping them when needed, providing them with suitable options and explaining the relevance of the tasks.

1.1.3 Mediators between teaching quality and student outcomes

Depth of processing. Based on cognitive constructivist learning theory (De Corte, 1995), depth of processing, or high-level thinking, is a student's reaction to cognitively activating teaching (Klieme & Rakoczy, 2008). The concept of depth of processing – the level at which a student processes what they are taught – encompasses critical thinking, reasoning, making sense, finding patterns, solving non-routine problems, as well as some aspects of self-regulation and metacognition (Baumert et al., 2010; Boston & Candela, 2018; Klieme et al., 2009; Lipowsky et al., 2009; Praetorius et al., 2018). Mathematics teaching, in particular, should incorporate challenging tasks that are neither too easy nor too hard so that students can develop an in-depth understanding of concepts, not just memorize facts (Hiebert & Grouws, 2007; Silver & Stein, 1996; Stein et al., 1996; Stein & Lane, 1996). Depth of processing has been empirically linked to student achievement (e.g., Chi & Wylie, 2014; Clifford, 1990; Lipowksy et al., 2009) and conceptual development (Stein et al., 1996; Stein & Lane, 1996). In the TBD model depth of processing mediates the relation between cognitive activation and student achievement, and classroom management is assumed to be directly related to depth of processing since a learning environment that helps students to pay attention is seen as an important prerequisite for in-depth engagement with a task (e.g., Lipowsky & Hess, 2019).

Time-on-task. Time-on-task is the class time during which students are actually engaged in learning activities contributing to learning gains and performance (Brophy, 2006; Emmer & Stough, 2001; Finn & Zimmer, 2012; Fisher et al., 1981; Rakoczy, 2006; Wang et al., 1993). In the TBD model, time-on-task is a response to classroom management, which in turn is a strong predictor of student learning and achievement (Böheim et al., 2020; Brophy, 2000; Hattie, 2009; Klieme et al., 2009; Seidel & Shavelson, 2007).

Need satisfaction. Research based on SDT resulted in the addition of the satisfaction of the three basic needs for autonomy, competence, and relatedness as a mediator between student support and motivation (Klieme & Rakoczy, 2003). The need for autonomy is the need to experience personal freedom, volition, and choice (Vansteenkiste et al., 2010). The need for competence is the student's desire for mastery and effectiveness during tasks (Ryan & Deci, 2002). The need for relatedness refers to the desire for close and warm relationships (Baumeister & Leary, 1995; Deci & Ryan, 2002). According to SDT teaching behaviors can influence whether student needs are satisfied (Black & Deci, 2000). Additionally, within the TBD model classroom management is an important prerequisite for the satisfaction of students' basic needs because, for example, well-organized, undisturbed classrooms may mean students feel more effective when performing tasks (Kunter et al., 2007).

1.2 Revisiting the TBD model

The TBD model assumes that classroom management has an influence on all three mediators (depth of thinking, time-on-task, and need satisfaction), and cognitive activation and student support affect depth of processing and need satisfaction respectively (see Figure 1). There is empirical evidence for the role played by single mediators (e.g., Hiebert & Grouws, 2007; Stein & Lane, 1996 for cognitive activation; Hospel & Galand, 2016; Kunter, et al., 2007 for classroom management; and Kiemer et al., 2018; Mouratidis et al., 2013 for student support). However, the current TBD model proposes a complex web of influences and theoretical assumptions which have been added incrementally over time. It is therefore important to periodically review and possibly revise these assumptions and the mediation paths proposed by Klieme et al. (2006). The need for a review has been underscored by recent evidence that several of the assumptions are not empirically supported (Praetorius et al., 2018). Therefore, robust model and theory building warrants a thorough revisit and in-depth investigation of the entire TBD model (Praetorius et al., 2020a).

Section 1.2.1 is a discussion of the possible alternate paths derived from established theories of motivation and cognition, such as expectancy-value theory (EVT; Wigfield & Eccles, 1992), that were not explicitly considered in the formulation of the TBD model but have considerable overlap with its core assumptions.

Relevant theories were systematically selected, by using the definitions of the dimensions and mediators within the TBD model and conducting a literature search for studies that assessed those constructs, including their subdimensions. Klieme et al. (2009) highlighted that while constructivist ideas play a crucial role in understanding teaching quality, they alone cannot fully explain the utilization of learning opportunities and the reasons behind such usage. Hence, the integration of motivational and cognitive theories seems essential to comprehensively grasp the learning processes involved. The objective was to improve the theoretical basis of the TBD model and provide a more comprehensive understanding of the underlying processes that affect how teaching quality impacts student outcomes.

When the theoretical views and their empirical insights were incorporated into the TBD model, it became evident that additional mediation paths may exist. For example, several theories in the domain of achievement motivation, such as interest theory (IT; Hidi & Renninger, 2006) and the control-value theory of achievement emotions (CVT; Pekrun, 2006), suggest that an optimal challenge or even being engaged in a task may affect not only achievement, but also motivational and cognitive processes (see for example Vu et al., 2022; Wentzel & Miele, 2016). We elaborate on these additional assumptions in the following.

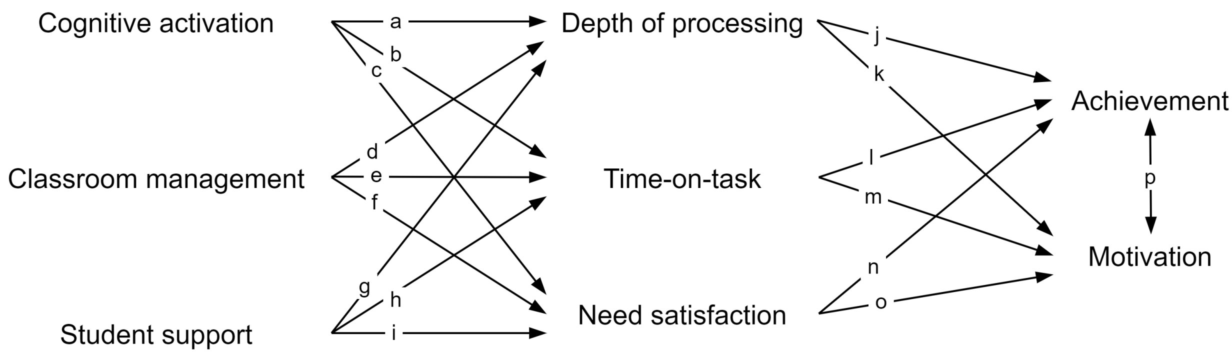

Figure 2. Possible relationships between the different parts of the TBD model.

1.2.1 Mediating paths for cognitive activation

The TBD model assumes a relation between cognitive activation and depth of processing (Figure 2, Path-a). However, theoretical and empirical evidence suggests that cognitive activation might also affect time-on-task (Figure 2, Path-b) and need satisfaction (Figure 2, Path-c). Cognitively challenging activities or tasks can direct student attention to particular aspects of content and specify methods by which information is processed and thus influence time-on-task (Doyle, 1983). This idea was also explored for mathematics teaching by Stein et al. (1996). For example, when a teacher asks questions or presents problems without obvious solutions, students are more likely to pay close attention. These arguments are consistent with influential views on achievement motivation. According to EVT student behavior can be seen as a product of the expectancy of success and value of reward (Atkinson, 1957; Heckhausen, 1991; Wigfield & Eccles, 1992).

The theory of motivational intensity (MIT) distinguishes between mere willingness to engage in a task and actual effort (Brehm & Self, 1989; Richter et al., 2016). According to this theory, conditions are identified which determine how much resource is allocated for engaging in a task. Moreover, a principle of resource conservation is proposed where it is assumed that even if the willingness to engage in a task is high, only as much effort as needed to succeed in a task will be allocated (Brehm & Self, 1989). If a task is very easy, effort will be low. When a task is too difficult or when the difficulty exceeds the value of a given reward, a student is likely to disengage from the task, resulting in diminished time-on-task. Given the fact that optimal task difficulty (Hiebert & Grouws, 2007), as well as adaptivity and individualization (Helm, 2016; Lotz, 2016; Rakoczy & Pauli, 2006) are often seen as parts of cognitive activation, an effect on time-on-task is also highly probable.

Theoretical and empirical evidence also suggests that cognitive activation can be related to students' satisfaction of basic psychological needs (Figure 2, Path-c). According to EVT (Wigfield & Eccles, 1992) and SDT (Ryan & Deci, 2017), when teachers give optimally challenging tasks, students' expectancies for success can be fostered (EVT; Wigfield & Eccles, 2002) and in a similar vein their competence need can be satisfied (SDT; Reeve, 2006; 2016). Similarly, the basic need for autonomy can be satisfied when teachers present non-routine problems, as it fosters students' critical thinking and encourages them to solve the tasks using their own methods, which is an important aspect of autonomy in the classroom (SDT; Reeve 2009; Reeve & Jang, 2006). If the students perceive the tasks as valuable and relevant, their basic psychological need for autonomy will be satisfied (SDT; Reeve & Jang, 2006). Empirical studies based on SDT support this link. For example, cognitive activation indirectly affected student interest and self-efficacy through autonomy and competence need satisfaction (Schukajlow et al., 2019; Schukajlow & Krug, 2014). Another study argued that a potential underlying mechanism between cognitive activation and student enjoyment in mathematics could be autonomy and competence need satisfaction (Lazarides & Buchholz, 2019). Moreover, cognitively activating behaviors such as aiming to foster independent engagement with the learning content, directly affect autonomy (Lotz, 2016). Since co-construction of knowledge is an important part of cognitive activation, the experience of relatedness could also be affected (see Ryan & Powelson, 1991; Sun & Chen, 2010 for the interplay and similarity of those constructs).

1.2.2 Mediating paths for classroom management

Within the TBD model *classroom management* is expected to affect all three mediators. Classroom management has been shown to affect time-on-task (Emmer & Stough, 2001; Finn & Zimmer, 2012; Fisher et al., 1981; Rakoczy, 2006; Wang et al., 1993) (Figure 2, Path-e), aspects of cognitive engagement (i.e., use of learning and self-regulation strategies) (Hospel & Galand, 2016) (Figure 2, Path-d), and students' need satisfaction (Kunter et al., 2007) (Figure 2, Path-f). Thus, classroom management should be relevant for all student learning processes (i.e., depth of processing, time-on-task, need satisfaction) in the classroom.

1.2.3 Mediating paths for student support

According to SDT, student support has a positive effect on the satisfaction of students' basic psychological needs (Ahn et al., 2021; Deci & Ryan, 2000; Jang et al., 2012; Kiemer et al., 2018; Mouratidis et al., 2013; Zhang et al., 2011) (Figure 2, Path-i). However, student support is also likely to be related to depth of processing (Figure 2, Path-g) and time-on-task (Figure 2, Path-h), which differs from what is postulated in the TBD model.

By engaging in supportive teaching behavior, characterized by mutual respect, teachers actively promote a positive learning environment. Students are not distracted by a negative teacher-student relationship that could elicit emotions that interfere with attention and self-regulation (Blair, 2002; Murray & Pianta, 2007). A good relationship between teachers and students also allows students to actively participate in their learning environment (Hughes et al., 2008; Pianta & Steinberg, 1992). Similarly, by giving constructive feedback, approaching student errors and misconceptions in a positive way, and monitoring student progress, teachers increase active learning time (Grabinger & Dunlap, 1995; Grabinger et al., 1997). Cognitive information processing theory (IPT; Atkinson & Shiffrin, 1968; Driscoll, 2005) states that students are attentive when they select and process information that is very important and meaningful for them. According to SDT, one key aspect of student support is making the information relevant and meaningful to the students (see also Ahmadi et al., 2023). For example, when teachers engage in autonomy supportive behaviors such as providing rationales for the content and personal relevance, then students are more likely to pay attention during the lesson because the information is useful, meaningful, and important to them (Lietaert et al., 2015).

A positive climate also allows students to try new and creative solutions without reservations (Chan & Yuen, 2014), an important aspect of depth of processing. This is because an encouraging, respectful, supportive, and positive learning environment that is open to creativity and improvement, encourages students to seek challenges (Turner & Meyer, 2004). In addition, according to the elaboration likelihood model (ELM, Petty & Cacioppo, 1986) when teachers highlight the relevance of the tasks to students, the students' personal involvement increases, which in turn fosters depth of processing (Illies & Reiter-Palmon, 2004; Petty et al., 1983; Mitchell, 1993).

Interest theory has been used to describe the relation between personal involvement, depth of processing, and time-on-task (Hidi & Renninger, 2006; Renninger & Hidi, 2002). When a student's attention is triggered by relevant tasks and personal involvement, they will also become interested in content. Several studies confirm the link between aspects of student support and aspects of depth of processing such as self-regulation and deep learning strategies (Hospel & Galand, 2016; Rieser et al., 2016; Ruiz-Alfonso & León, 2019; Wang & Eccles, 2013), higher analytical problem-solving skills, and student challenge preferences (Boggiano et al., 1988; 1993; Guay et al., 2008). Positive relations have also been identified between student support and time-on-task (Chiu, 2004; Deci et al., 1994; Stallings, 1980). All these studies lend weight to the hypothesis that student support can predict depth of processing and time-on-task.

1.2.4 Student use of opportunities and student outcomes

Because the mediators are interrelated, the relationship between the mediators and the outcomes might also be less discrete than how they are shown in the original model (Figure 1); the original model already indicated the relationship between motivational outcomes and achievement (Figure 2, Path-p). Other important theories, such as CVT (Pekrun, 2006), also suggest that depth of processing and time-on-task could be related to motivational outcomes (Figure 2, Paths k and m). For example, students who think critically and solve modelling problems by constructing multiple solutions have a greater interest in the subject (Schukajlow & Krug, 2014) and higher self-efficacy (Schukajlow et al., 2019). Interest and self-efficacy have been treated as motivational outcomes in TBD research (Figure 2, Path-k) (Dorfner et al., 2018; Fauth et al., 2014, 2019; Förtsch et al., 2017; Li et al., 2020).

Time-on-task not only promotes academic achievement (Evertson & Harris, 1992; Good & Brophy, 2003), but also appears to be relevant for fostering student motivation (Butler & Shibaz, 2008; Lazarides & Buchholz, 2019; Rakoczy, 2006) (Figure 2, Path-m). This relationship is also suggested by other motivation theories such as interest theory and CVT. In these instances, it is hypothesized that being on task or processing information at a deep level creates positive emotions for students (i.e., activity emotions), which in turn fosters their interest and motivation.

Finally, as proposed in the TBD model, it is hypothesized that need satisfaction affects motivational outcomes which in turn affect achievement (Figure 2, Path-o-p). Studies have shown a link between the satisfaction of a student's needs and their autonomous motivation (e.g., Mouratidis et al., 2015; Ryan & Deci, 2009), interest (e.g., Kunter et al., 2007), and self-efficacy (e.g., Sun et al., 2020; Zhen et al., 2017) (Figure 2, Path-o). However, according to SDT, when students' basic psychological needs are satisfied, they display improved academic performance and achievement (Ryan & Deci, 2017). Theoretical considerations based on SDT, in combination with the studies which found positive relationships between need satisfaction and student achievement (Badri et al., 2014; Wang et al., 2019; Zhou et al., 2021), lead us to hypothesize that need satisfaction is directly positively related not only to motivational outcomes, but also to achievement. Depth of processing and time-on-task are likely to be linked to motivational outcomes (Figure 2, Paths k and m) and need satisfaction can be directly related to achievement (Figure 2, Path-n).

1.3 Study

A review of theories in the field of cognitive and motivational psychology and related empirical evidence strongly suggests that there should be more mediation paths than those which have been discussed in the TBD literature to date. Our assumptions will be tested by constructing models which consider the assumptions of the original TBD model and additional possible paths. Our concrete hypotheses are as follows:

H1: The three basic dimensions of teaching quality are all related to the development of student achievement and interest.

H2: Cognitive activation indirectly predicts the development of student achievement and interest through depth of processing, time-on-task, and need satisfaction.

H3: Classroom management indirectly predicts the development of student achievement and interest through depth of processing, time-on-task, and need satisfaction.

H4: Student support indirectly predicts the development of student achievement and interest through depth of processing, time-on-task, and need satisfaction.

2. Method

This study investigates whether student perceptions of cognitive activation, classroom management, and student support indirectly affect student achievement and interest in mathematics through depth of processing, time-on-task, and need satisfaction. We analyzed data collected in Germany as a part of the Teaching and Learning International Survey (TALIS) Video Study conducted by the Organisation for Economic Co-operation and Development (OECD, 2020).

2.1 Participants and procedures

The study sample was selected from participants in the TALIS Video Study for Germany using convenience sampling. The initial sample consisted of 1143 students from 50 classrooms and 39 schools. There are big differences in learning goals, school curricula, student achievement levels, class composition and individual student characteristics between school tracks in Germany (Hachfeld & Lazarides, 2020). As most participating classrooms were from the academic track ("Gymnasium") and we were interested in a homogeneous sample so that the data can be interpreted more unambiguously, we removed participants from all other school forms (e.g., lower-track schools and vocational schools). The final sample consists of 958 students from 41 classrooms and 30 schools (M_{age} = 14.82, *SD* = 0.62; 50.5 % females; 5.3% did not report their gender). The average number of students per classroom was 23.37 (*SD* = 4.73, min = 11 max = 31). Of 41 classrooms, the majority (35) were 9th grade level and six were 8th grade. Most of the students reported that they were born in Germany (*n* = 869), *n* = 36 students reported that they were born in other countries, and *n* = 53 did not report their country of birth.

The TALIS Video Study conformed to ethical standards (OECD, 2020). School principals, teachers, students, and their parents were informed about the purpose of the study. The participants were assured that their participation was anonymous and voluntary and that their information would be secure and confidential.

2.2 Instruments and measures

The student survey asked about family and peer circumstances and aspects of students' cognitive, motivational, and emotional learning. It also asked students for their perceptions of teaching quality in the mathematics lessons at the beginning of a specific teaching unit, *quadratic equations* (McCaffrey et al., 2020; Praetorius et al., 2020b). In the TALIS Video Study, the constructs measured in the pre-test (T1) are operationalized in terms of mathematics in general, whereas the constructs measured in the post-test (T2) are operationalized only in terms of quadratic equations. To test our hypotheses, variables from the first and second measurement points were used. The unit included between 360 and 1080 minutes of lesson time (*M* = 797.20, *SD* = 173.59), spread over a period of 22 to 130 days (*M* = 58.83, *SD* = 26.01) (see Supplementary Material).

Teaching quality dimensions were measured using student rating, which is considered a valid, reliable, and efficient measure of teaching quality (van der Scheer et al., 2019). Depth of processing, time-on-task, need satisfaction, and interest were assessed using student self-reports. Self-reports are useful for assessing constructs that are not directly observable, such as student use of learning opportunities (Appleton et al., 2006; Fredricks & McColskey, 2012).

Many items in the TALIS Video Study questionnaire were based on previous TALIS and Programme for International Student Assessment (PISA) studies (OECD, 2020; Praetorius et al., 2020 b). The concrete item wordings of the assessed constructs are shown in Appendix A. Each item was assessed using a four-point Likert scale. Negative items in the questionnaire were reverse-coded. To account for level-specific reliability (Geldhof et al., 2014), we calculated McDonald's omega (ω; McDonald, 1999) for both the within and between levels; these are reported in Table 1. We also calculated the descriptive statistics for each item and each subscale (see Supplementary Material).

2.2.1 Independent variables: Three dimensions of teaching quality (TBD)

In the TALIS Video Study for Germany, student perceptions of cognitive activation, classroom management, and student support were assessed using items similar to those used in previous TALIS and PISA studies (OECD, 2020). Student-reported cognitive activation was assessed with seven items designed to reveal their perceptions of whether teachers presented tasks and their solutions in a manner that would promote conceptual understanding and content-based discourse (e.g., "Our mathematics teacher gives tasks that require us to think critically"). Student-reported classroom management was initially assessed using 10 items related to disruptions, transitions, monitoring, and clarity of rules (e.g., "In the lesson, our teacher is clear to us why certain rules are important"). However, we excluded two items from the study because of negative correlations between them and other classroom management items, resulting in eight items for measuring classroom management (see Table 1). Student-reported student support was assessed using 11 items including three items covering teacher support, four items covering autonomy support, and four items covering competence support (e.g., "Our mathematics teacher makes me feel confident in my ability to learn the material").

2.2.2 Mediators: Student use of learning opportunities

Student use of learning opportunities was assessed using three scales (OECD, 2020; Vieluf et al., 2020). Student self-reported depth of processing was assessed with three items (e.g., "I keep thinking about tasks until I really understand them"). Student self-reported time-on-task was assessed using three items (e.g., "I pay attention in mathematics class"). Student self-reported need satisfaction was assessed using three items (e.g., "I feel I can decide on things on my own"). In designing our analyses, we had to decide on whether using learning processes at either T1 or T2 as mediators. The primary goal of this study was to examine whether the longitudinal effects of teaching quality on student outcomes were mediated by learning processes. Teaching quality was measured focusing on teaching in a class in general (T1) whereas the outcomes were focusing on the unit on quadratic equations (T2). In choosing the appropriate time point for the mediators, we therefore had to decide to either assess learning processes closer to the outcomes or closer to teaching quality. We opted for the latter due to the close interplay between opportunities and use.

2.2.3 Outcomes: Interest and Achievement

We chose student individual interest in mathematics classes, which was also used as an outcome in the report of TALIS Video Study and by other studies, as a motivational outcome (Herbert et al., 2022; Zhu & Kaiser, 2022). Student self-reported interest in mathematics classes was assessed at T1 using three items (e.g., "I often think that what we are talking about in my mathematics class is interesting."). Student self-reported interest in the instructional unit was assessed after the unit, at T2, using three items (e.g., "I was interested in the topic of quadratic equations").

Students' knowledge of mathematics was assessed using 30 multiple-choice items. The pre-test focused on the key prerequisites for the conceptual understanding of quadratic equations. Items covered students' precursors to understanding quadratic equations such as numbers, algebraic expressions, and algebraic equations. The post-test (T2) focused on students' knowledge of quadratic equations and its applications (McCaffrey et al., 2020).

2.3 Data analysis

We tested our hypothesized mediation paths with correlative (preliminary) and longitudinal (main) analyses using the lavaan package (v0.6-8; Rosseel, 2012) in the R programming software (R Development Core Team, 2020). The R code for all the analyses is included in the Supplementary Material.

To assess the reliability of the aggregated student variables, intraclass correlation coefficients (ICC1 and ICC2) were computed for all model variables (see Table 1). ICC1 ranged between 4% and 37%. This range shows the extent to which the individual ratings of the variables are attributable to classroom membership (LeBreton & Senter, 2008). ICC2 is the reliability of the class-average constructs and ranged between .45 and .93. ICC2 values between .70 and .85 indicate acceptable levels of reliability (LeBreton & Senter, 2008; Lüdtke et al., 2009).

To account for the hierarchical structure of the data, the main analyses were multilevel longitudinal path analyses. Due to the complexity of the TBD model, we investigated the mediating effects of three mediators for each dimension of teaching quality in separate models. In keeping with the methodology employed in other empirical studies investigating mediation between teaching quality and student achievement (e.g., León et al., 2017; Ruiz Alfonso & León, 2017; Theis et al., 2020), we used 1-1-1 and 2-2-2 models so that between groups effects and within-group effects were separated (Preacher et al., 2011). Because the cluster size was too small to apply latent models and due to the complexity of the models, we averaged the items per scale and used the resulting mean scores as manifest variables in our path models. Moreover, due to the non-normality of the assessed variables, we used the maximum likelihood with robust standard errors (MLR) estimator (Savalei & Rosseel, 2021).

Fifty-one students for whom all values were missing for all the assessed variables were removed from the analyses, leaving a total sample of $n = 907$. The percentage of missing values for the assessed scales in the total sample ranged from 0.1% to 0.7%. We did not apply a special missing value treatment because of this low percentage (Kline, 2011).

We used the comparative fit index (CFI), the Tucker-Lewis index (TLI), the standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA) to evaluate model fit. Adequate and very good fit is achieved when the CFI and TLI are greater than .90 and .95 respectively; RMSEA and SRMR show adequate fit when they are between .05 and .08, and they show very good fit when they are less than .05 (Hu & Bentler, 1999). Because distributions of indirect effects could be non-normal, the bootstrapping method was used to calculate confidence intervals for the indirect effects ($N = 1000$ bootstrap samples; Preacher & Hayes, 2008). The indirect effects are considered statistically significant when the 95% confidence intervals do not include zero (Cheung & Lau, 2008; Mackinnon et al., 2004).

3. Results

3.1 Descriptive statistics and bivariate correlations

Descriptive statistics and bivariate Pearson's correlations for all observed variables at the classroom and student level are presented in Table 1. Positive correlations between the independent variables (the three basic dimensions of teaching quality) and all the mediators (depth of processing, time-on-task, and need satisfaction) were found at both classroom and student levels. However, not all the expected correlations between independent variables and outcomes (interest and achievement), as well as between the mediators and outcomes were found. For example, at the student level, the three basic dimensions of teaching quality were not correlated with achievement at T1.

3.2 Preliminary analyses

3.2.1 Relationships between teaching quality, learning processes, and student outcomes

As a first step, we conducted multilevel path analyses using all the variables that had been assessed at the same point in time (i.e., T1). The results of the three separate direct effect models indicated that, at the student level, all the three basic dimensions of teaching quality were positively related to student interest and only student support was positively related to student achievement. At the classroom level, classroom management and student support were positively related to student interest, and classroom management was positively related to achievement.

In a second step, we tested the three mediation models. The correlational mediation analyses indicated that all the three basic dimensions were positively related to all the mediators at both student and classroom levels. Furthermore, positive associations were found between all the mediators and outcomes at both levels, except for time-on-task and achievement (for details, see Supplementary Material).

3.3 Main analyses

3.3.1 Longitudinal relationships between teaching quality and student outcomes

We estimated three multilevel longitudinal path analyses. By conducting three separate direct effect models using longitudinal data, we tested the direct relationship between the three dimensions of teaching quality at T1 and student achievement and interest at T2 while controlling for student achievement and interest, respectively, at T1. The model fit indices are sufficient except for TLI, which is slightly lower than the acceptable values for two of the three models (see Table 2). The results of the three multilevel path models indicate that the three basic dimensions of teaching quality were not directly associated with mathematics achievement and interest at T2, neither at the classroom nor at the student level, controlling for student achievement and interest at T1 (see Figure 3). Strong positive relationships between T1 interest and T2 interest and between T1 achievement and T2 achievement were found at both the classroom and student levels. At the student level, the three dimensions of teaching quality at T1 were positively associated with student interest at T1, whereas at the classroom level, only student support at T1 was found to be positively related to student interest at T1.

Table 1

Descriptives, ICCs, reliability estimates (ω), and within and between level intercorrelations between the measured variables

	1	2	3	4	5	6	7	8	9	10
1. Cognitive activation (T1)	1	.15***	.48***	.26***	.17***	.32***	.27***	-.02	.16***	-.01
2. Classroom management (T1)	.14***	1	.31***	.07	.18***	.22***	.16***	-.02	.12**	.00
3. Student support (T1)	.49***	.22***	1	.32***	.26***	.60***	.46***	.08	.28***	.10
4. Depth of processing (T1)	.24***	.09	.32***	1	.34***	.45***	.54***	.25***	.36***	.25***
5. Time-on-task (T1)	.14***	.21***	.21***	.34***	1	.26***	.33***	.07	.27***	.06
6. Need satisfaction (T1)	.33***	.20***	.62***	.47***	.26***	1	.52***	.25***	.28***	.21***
7. Interest (T1)	.24***	.18***	.48***	.54***	.35***	.54***	1	.22***	.56***	.25***
8. Achievement (T1)	-.03	.08	.03	.27***	.10*	.22***	.18***	1	.09	.58***
9. Interest (T2)	.15***	.12**	.30***	.38***	.28***	.30***	.58***	.06	1	.18***
10. Achievement (T2)	-.05	.10*	.08	.27***	.08	.20***	.24***	.61***	.21***	1
$\text{Mean}_{\text{within}}$	2.62	3.01	2.98	2.66	3.10	2.86	2.40	.73	2.17	.50
SD_{within}	.49	.47	.56	.63	.54	.64	.78	.15	.73	.19
ICC1	.10	.37	.26	.04	.04	.11	.13	.20	.12	.16
ICC2	.72	.93	.89	.50	.45	.73	.77	.85	.73	.80
ω_{within}	.65	.60	.86	.67	.73	.63	.85	-	.81	-
ω_{between}	.74	.95	.98	1.00	.95	.90	.98	-	.99	-

Note. *$p < .05$. **$p < .01$. ***$p < .001$. Student-level correlations are displayed below the diagonal and classroom-level correlations are displayed above the diagonal.

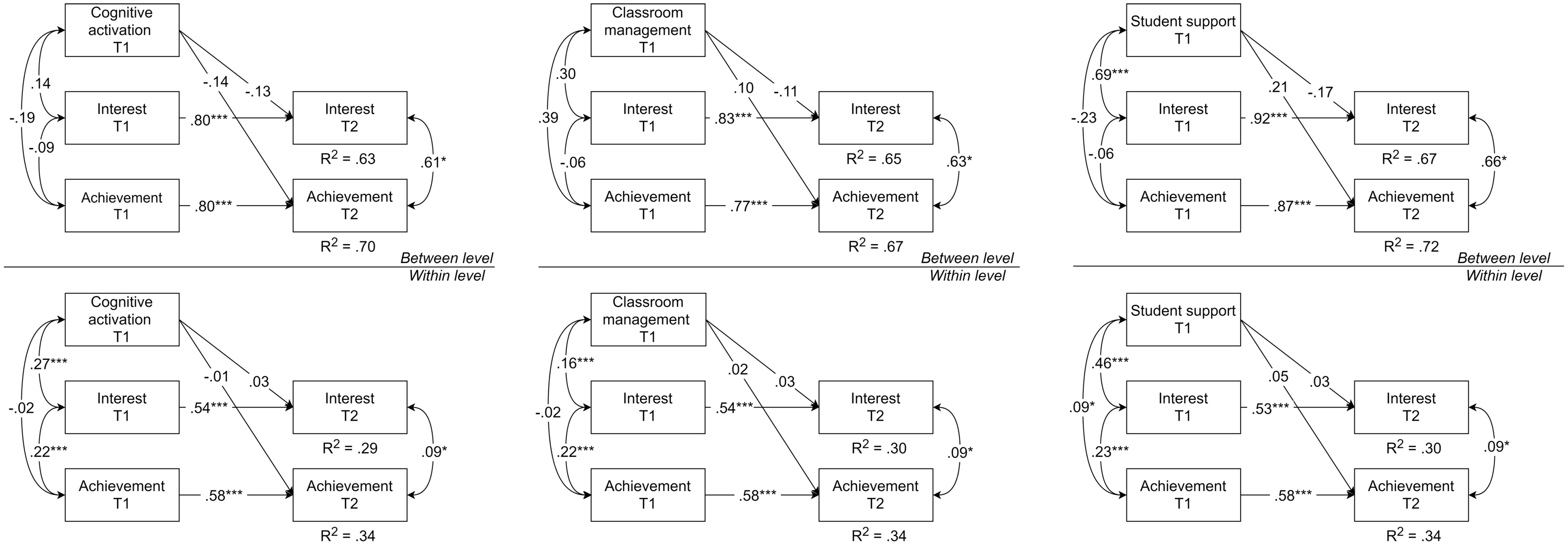

Figure 3. Models of the effects of three basic dimensions of teaching quality at T1 on student achievement and interest at T2 controlling for student achievement and interest at T1.

Note. *$p < .05$. **$p < .01$. ***$p < .001$.

Table 2

Model fit indices

Model fit indices	$\chi2$	*p*	CFI	TLI	RMSEA [90%-CI]	$SRMR_{within}$	$SRMR_{between}$
Direct effects models							
1. Cognitive activation	34.841	< .000	.958	.790	.098 [.067 - .132]	.037	.073
2. Classroom management	21.900	< .000	.977	.885	.075 [.047 - .106]	.036	.055
3. Student support	17.571	= .001	.985	.924	.065 [.037 - .096]	.030	.024
Mediation models							
4. Cognitive activation	9.294	= .054	.997	.952	.041 [.000 - .075]	.013	.026
5. Classroom management	7.624	= .106	.998	.967	.034 [.000 - .068]	.012	.022
6. Student support	8.557	= .073	.997	.964	.038 [.000 - .072]	.012	.015

3.3.2 The mediating role of student use of learning opportunities

After investigating the direct effect models, we constructed three separate multilevel longitudinal mediation models for each of the three dimensions of teaching quality: cognitive activation, classroom management, and student support.

As with the direct effect models, strong positive relationships between interest at T1 and at T2 and between achievement at T1 and T2 were found at both the classroom and student level.

All three mediators – depth of processing, time-on-task, and need satisfaction – were added to the direct effect models. The multilevel mediation models yielded satisfying fit indices (see Table 2). We also conducted bootstrap analyses to calculate the mediating effects of depth of processing, time-on-task, and need satisfaction because the distribution of the mediating effects can be non-normal (Preacher & Hayes, 2008). The results of the bootstrap analyses are shown in Table 3. The three multilevel mediation models generated the following results:

Cognitive activation: Cognitive activation at T1 was positively related to need satisfaction at T1 at both the student and classroom level, and depth of processing at T1 at the student level (see Figure 4). Mediation analyses revealed that depth of processing at T1 mediated the relation between T1 cognitive activation and T2 achievement ($\beta = .02$; $B = .01$; $SEB = .00$; 95%-CI: [.00, .01]) at the student level.

Classroom management: Classroom management at T1 was positively related to time-on-task at T1 at both the student and the classroom level, and need satisfaction at T1 at the student level (Figure 5). Mediation analyses also confirmed the mediating role of T1 time-on-task between T1 classroom management and T2 interest ($\beta = .01$; $B = .02$; $SEB = .01$; 95%-CI: [.00, .04])

Student support: For student support, a positive relation between T1 student support and T1 need satisfaction was found both at the classroom level and at the student level (see Figure 6). Relations with T1 depth of processing and T1 time-on-task were also found at the student level. Mediation analyses showed that T1 depth of processing mediated the relation between T1 student support and T2 achievement ($\beta = .01$; $B = .00$; $SEB = .00$; 95%-CI: [.00, .01]), whereas T1 time-on-task mediated the relation between T1 student support and T2 interest ($\beta = .01$; $B = .02$; $SEB = .01$; 95%-CI: [.00, .03]) at the student level.

It is important to note that some of the standardized regression (beta) coefficients in the models are larger than 1.00. This is primarily due to the multicollinearity and low variance of the variables at the classroom level (i.e., ICCs of depth of processing and time-on-task were .04).

In summary, in our analyses, none of the mediation assumptions were supported at the classroom level, but some of them were supported at the student level.

Table 3

Results of bootstrap analyses

Mediation Models	Classroom level				Student level			
Model 4. Cognitive activation →	*β*	*B*	SE_B	%95CIs	*β*	*B*	SE_B	%95CIs
Depth of processing → achievement	.00	.00	.02	[-.03, .03]	.02	.01	.00	[.00, .01]
Depth of processing → interest	.00	.00	.01	[-.02, .02]	.01	.02	.01	[-.00, .04]
Time-on-task → achievement	-.10	-.04	.06	[-.15, .07]	-.00	-.00	.00	[-.00, .00]
Time-on-task → interest	.02	.02	.11	[-.19, .23]	.01	.01	.01	[.00, .02]
Need satisfaction → achievement	-.02	-.01	.08	[-.17, .16]	.01	.00	.00	[.00, .01]
Need satisfaction → interest	-.13	-.19	.33	[-.84, .45]	-.01	-.01	.01	[-.03, .01]
Model 5. Classroom management →								
Depth of processing → achievement	-.01	-.00	.01	[-.02, .02]	.00	.00	.00	[-.00, .00]
Depth of processing → interest	-.00	-.00	.02	[-.03, .03]	.00	.00	.01	[-.01, .01]
Time-on-task → achievement	.30	.07	.06	[-.04, .19]	-.00	-.00	.00	[-.01, .00]
Time-on-task → interest	.05	.04	.18	[-.32, .39]	.01	.02	.01	[.00, .04]
Need satisfaction → achievement	.00	.00	.00	[-.01, .01]	.01	.00	.00	[-.00, .01]
Need satisfaction → interest	.01	.01	.04	[-.07, .09]	-.01	-.01	.01	[-.03, .01]
Model 6. Student support →								
Depth of processing → achievement	-.00	-.00	.02	[-.04, .04]	.01	.00	.00	[.00, .01]
Depth of processing → interest	.00	.00	.01	[-.02, .02]	.01	.01	.01	[-.00, .03]
Time-on-task → achievement	.01	.00	.04	[-.08, .08]	-.01	-.00	.00	[-.01, .00]
Time-on-task → interest	.03	.03	.10	[-.18, .23]	.01	.02	.01	[.00, .03]
Need satisfaction → achievement	-1.17	-.28	.20	[-.67, .11]	.02	.01	.01	[-.00, .02]
Need satisfaction → interest	-.29	-.23	.68	[-1.56, 1.10]	-.02	-.02	.02	[-.07, .02]

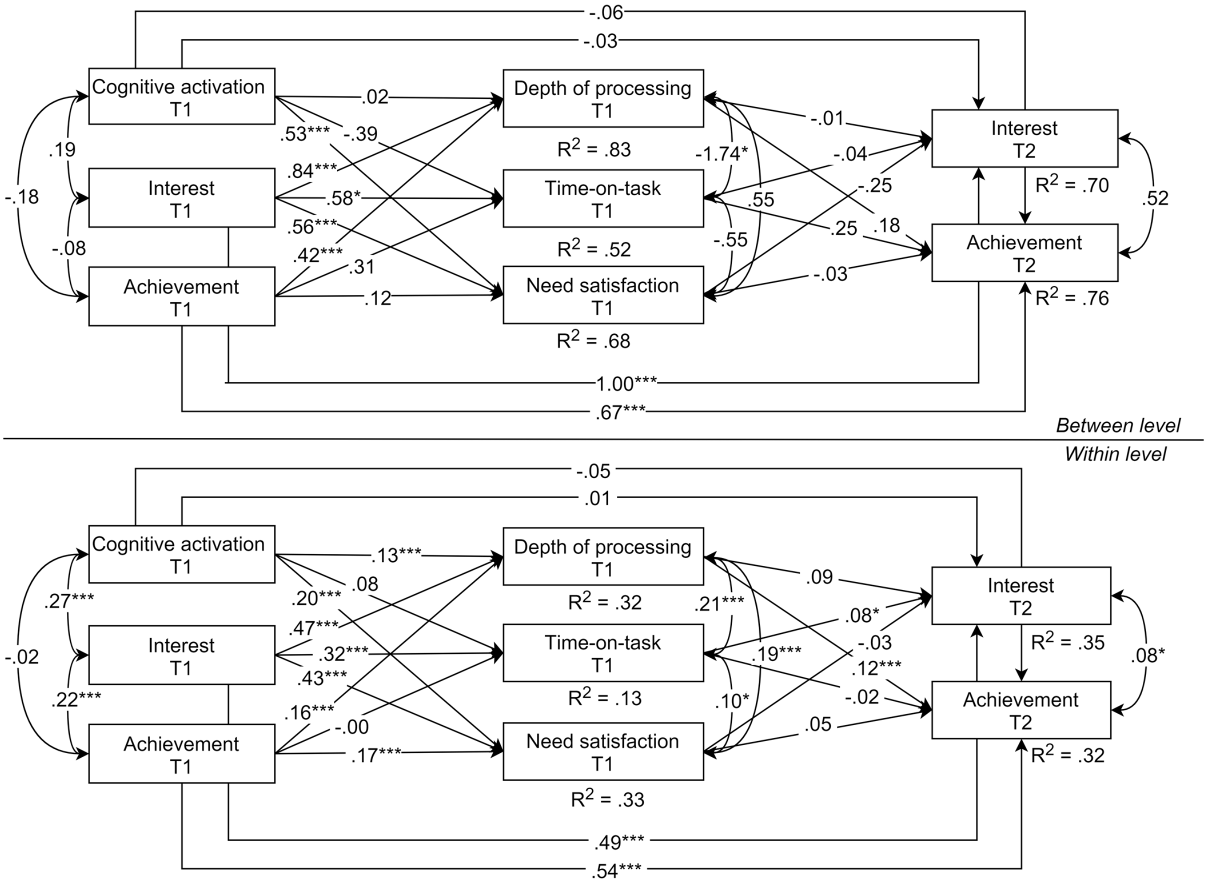

Figure 4. Mediation model for cognitive activation with standardized coefficients.

Note. **p* < .05. ***p* < .01. ****p* < .001.

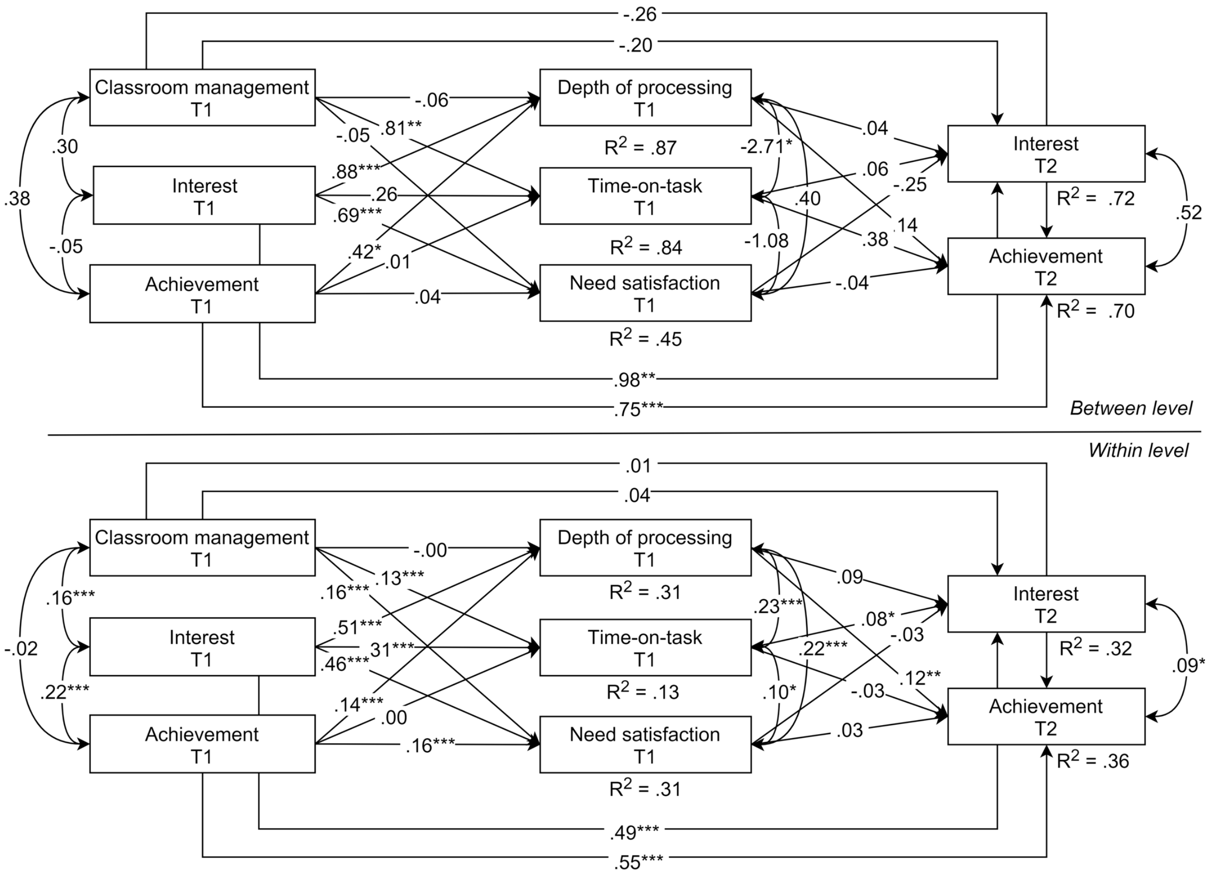

Figure 5. Mediation model for classroom management with standardized coefficients.

Note. *$p < .05$. **$p < .01$. ***$p < .001$.

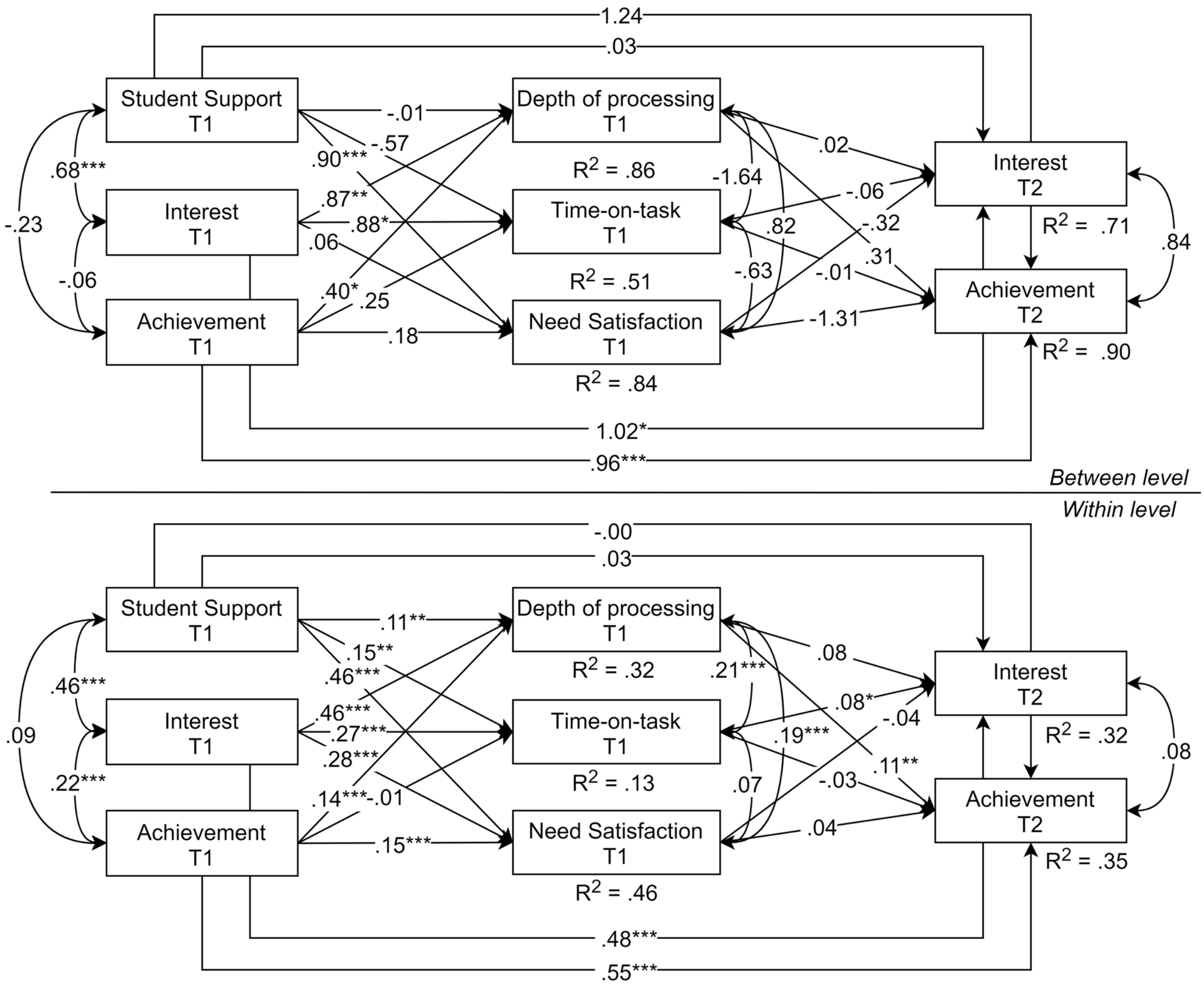

Figure 6. Mediation model for student support with standardized coefficients.

Note. **p* < .05. ***p* < .01. ****p* < .001.

4. Discussion

This study aimed to investigate the mediating role of student learning processes in the relationship between the three basic dimensions of teaching quality and student outcomes in mathematics by focusing on and extending on the hypotheses of the TBD model (Klieme et al., 2009). Contrary to the premise of the TBD model and our reasoning, the results of our study showed no statistically significant direct longitudinal effects of the teaching quality dimensions on student outcomes at either the classroom or individual level. Positive associations were found between teaching quality dimensions and mediators, but the partial mediation models at the classroom level failed to confirm the mediation hypotheses. At the student level, consistent with the TBD model, depth of processing mediated the relationship between cognitive activation and achievement. However, contrary to the predictions of the TBD model and consistent with our new hypotheses, the following relationships were found at the student level: Time-on-task mediated the relationship between classroom management and interest, and between student support and interest. Depth of processing mediated the relationship between student support and achievement. These results suggest that the TBD model could benefit from an expansion of its hypotheses about mediators using relevant theoretical approaches such as EVT (Wigfield & Eccles, 1992) and ELM (Petty & Cacioppo, 1983).

4.1 Conceptual expansion of the TBD model

The varied results of this study suggest that the relationships between the variables are more complicated than we had initially predicted. In its current form, the TBD model is based on the assumption that there is a clearly defined sequence of teaching quality dimensions over their associated mediators on student outcomes. The simple structure of the model was possibly deliberate, but it has resulted in a model that struggles to reflect the full complexity of teacher-student interactions (Vieluf & Klieme, 2023). It is therefore important that the model is expanded and refined. Based on our results, we believe that subsequent research needs to reassess three key assumptions of the current model:

The first fundamental assumption of the TBD model is that the variables are related in a specific, predefined manner, with quality dimensions preceding mediators in the model's structure. However, within the dimensions and mediators, there is no hierarchical distinction, implying equivalence among the entities within those categories. Our study found that not all the hypothesized relationships we thought might exist between the mediators (depth of processing, time-on-task, and need satisfaction) and achievement and interest outcomes were supported. Specifically, at the student level, we found that time-on-task predicted student interest and depth of processing predicted student achievement, but none of the other postulated relationships were observed. These findings could suggest that the mediators are not working similarly and in parallel. For instance, according to SDT (Ryan & Deci, 2017) need satisfaction may be a pre-condition for time-on-task and depth of processing, Schlesinger and Jentsch (2016) argued that time-on-task might be necessary for deep processing to occur, and according to Brown and Ryan (2003), conscious attention might be needed to meet psychological needs. We recommend that future research addresses the relationship between the three mediators such as considering the possibility that the mediators are sequential; one acts as a pre-condition for another. Similarly, there might be a hierarchical sequence also for teaching quality dimensions. For example, we were not able to confirm our assumption on the relation between classroom management and depth of processing at both student and class level. One explanation could be that classroom management only indirectly influences depth of processing through cognitive activation (Charalambous & Praetorius, 2020; Klieme et al., 2001). It acts as a pre-condition for the other teaching quality dimensions. This idea is supported by empirical evidence that classroom management predicts cognitive activation at the classroom level (Dorfner et al., 2018). Classroom management alone may not be enough to promote depth of processing, but it may play an enabling role. Researchers should continue to explore the existence of potential mediators between classroom management and depth of processing.

The second assumption of the current TBD model that our findings cast doubt on is that only three aspects of learning processes mediate the relationship between the three basic dimensions of teaching quality and student outcomes. Being based on the TBD, our paper focused on an analysis of these aspects. However, given the various theoretical approaches explored in this study, other aspects such as emotions related to achievement (CVT; Pekrun, 2006) and expectancies and values (EVT; Wigfield & Eccles, 1992) could also be included in the model. The inclusion of these mediators in particular could be productive since studies supporting the relationship between the three basic dimensions of teaching quality and these learning processes already exist (e.g., Burić & Kim, 2020; Lazarides & Buchholz, 2019).

The third assumption of the TBD model is that some variables are related to others in only one direction. We investigated the relationship between teaching quality at T1 and mediators at T1 and the relationship between mediators at T1 and achievement and interest at T2. We found depth of processing at T1 predicted student interest at T2. But this relationship might be bi-directional over the long term (Hidi & Renninger, 2006), i.e., when students are interested in mathematics classes, they tend to think more critically and try to solve more challenging problems. Likewise, our study showed that interest and achievement at T1 are related to mediators at T1 but the direction of the effects could not be ascertained because they were investigated at the same point in time. This is also true for the relations between teaching quality dimensions and mediators. Therefore, future studies should consider using more suitable designs such as cross lagged models and three measurement points to separately investigate the longitudinal mediating effects of each mediator.

Moreover, the relationships may also be influenced by control variables and moderators, such as student personality traits, adding yet more complexity to any analysis. Study designs should test and expand theoretical assumptions, using robust experimental or intervention designs. It is also vital to acknowledge the complexity of an educational reality encompassing countless interactions between teachers and students, not unfairly described as a “hall of mirrors” (Berliner, 2002, Cronbach, 1975). It is essential to recognize that continued exclusive reliance on quantitative methods such as mediation analysis may not capture the full complexity of the system. Qualitative approaches and mixed method studies are needed to further develop the TBD model (Vieluf & Klieme, 2023).

4.2 Correlational vs. longitudinal evidence

Our study highlighted that choosing whether to use correlational or longitudinal analyses can have a significant impact on the results. The direct effect models using a correlational design resulted in mostly positive direct associations. Contrary to our hypotheses, direct effect models with a longitudinal design revealed that at both levels, cognitive activation, classroom management, and student support did not directly predict achievement or interest. In correlational mediation models all paths, with the exception of the relationship between time-on-task and achievement, showed positive associations at both levels. However, in longitudinal mediation models, mediating effects were found only at the student level and only few of them could be identified.

Correlational research design is frequently used to confirm theoretically predicted relationships between variables in educational research because it is a practical approach. Most of the relationships we found using a correlational design were positive but the same was not true when the data were analyzed longitudinally. This discrepancy is important and researchers should investigate differences between correlational and longitudinal data in other settings.

Correlational results do not establish causality or the direction of effects. To avoid potential misconceptions, researchers should not rely only on correlational designs for research that may have practical implications for teachers. Also, the interpretation of correlational findings needs careful framing. For instance, correlational studies should avoid using directional language such as “affect” or “predict” to minimize potential misinterpretations. Although correlational studies can be a practical tool in the early stages of a new area of research, helping to identify any relationships, when the research field is saturated with the correlational studies, as it is in teaching quality research, we recommend the use of stronger methods such as longitudinal or experimental designs so that the directionality of effects can be established.

Although less used, longitudinal designs have the advantage of being able to reveal the direction of effects. They do, however, pose challenges. Firstly, using short time intervals between measurement points in longitudinal studies often results in a high stability of the variables over time (Begrich et al., 2023). This issue was observed in the analysis of student achievement and interest in our study. The high stability of outcome variables implies that the remaining variables, such as the dimensions of teaching quality at T1 or mediators at T1, only explain a little of the variance (Adachi & Willoughby, 2015; Praetorius et al., 2018; Warner et al., 2017). In the future, researchers could mitigate this effect by having longer intervals between measurement points. Extended intervals would also enable the monitoring of significant transitions, such as a change of teacher or shifts in classroom dynamics (Begrich et al., 2023). Another interesting avenue for future research would be to examine whether study outcomes are affected by time between measurements. In our study there was considerable variance in intervals, from 22 to 130 days. It would be interesting to analyze the differences between classes where the interval was larger and those where it was smaller by for example, dividing data at the median time interval, to determine the effect of time intervals on the stability of outcomes. However, due to the limitations imposed by the relatively small size of our study sample and the limited number of classrooms, it was not possible to run such a complex model (Hox & McNeish, 2020; Maas & Hox, 2005). Secondly, despite providing valuable insights, longitudinal studies only assess specific time points and do not denote any causal link between variables. Therefore, experiments or interventions are required to confirm the effects between the variables or determine the absence of effects in certain contexts and settings. For example, teaching quality could be manipulated by training a group of teachers to set optimally challenging tasks that cater to the level of each student. The results from this group could then be compared to a control group providing regular lessons using an experience sampling approach to investigate what effect the treatment had on their learning processes and outcomes (see Schukajlow et al., 2023; Talić et al., 2022).

While correlational studies give an initial indication of the relationships between variables, sometimes, these relationships are not confirmed by a longitudinal study, as is the case here. Longitudinal design in TBD research can be improved by having longer intervals between measurement points. However, more rigorous and holistic approaches are necessary in order to be able to show the effect of teaching quality on learning processes and then, in turn, on student outcomes.

4.3 Level of analyses

We ran the models at both the student and classroom levels and the results differed, depending on the level of analysis. In order to better understand how much individual student perceptions differed from the shared class perception, we separated within group and between group effects (Fauth et al., 2014; Marsh et al, 2012). This was also helpful for identifying the most suitable constructs for each level. For example, at the classroom level the low ICC1 of depth of processing and time-on-task showed that only 4% of the variance in those variables could be attributed to classroom membership. These two variables are also problematic regarding their low ICC2. Although considering between-level effects for variables with low ICCs is possible when intraclass correlations are nonzero and number of individuals per group is high (Julian, 2001; Lazarides & Buchholz, 2019), these effects must be interpreted with caution.

Given the low ICCs, especially for depth of processing and time-on-task, it appears that these constructs might be more idiosyncratic. While students in the same classroom are taught by a single teacher and may share some learning processes related to their common activities, such as solving specific mathematical problems (e.g., Hill & Rowe, 1996), each student-teacher interaction remains unique. This is because each student has different personality traits, beliefs, values, and a different ability level, prior knowledge, and family background (Helmke, 2012; Seidel, 2014), all of which will probably influence how they perceive any activity or teaching approach. This observation has important implications. Although researchers have been mostly treating teaching quality as a classroom level construct, it might be more important to consider teaching at both levels, paying attention to the individual level effects. Studies which mostly assessed teaching quality at classroom level could have missed the effect of individual variables.

A study can be interested in relations at the classroom level, the student level, or both (Senden et al., 2023; Stapleton et al., 2016). While the levels of analysis in a study depend on the data and research questions (Marsh et al., 2012), teaching and learning occur at both student and classroom levels and the effects at each level might be different. Although studies consider teaching quality most often at the classroom level, it is important that we do not ignore the effect of student-perceived teaching quality at the individual level. Researchers should consider refining operationalizations of teaching quality to include aspects such as differentiation and adaptivity (Vieluf & Klieme, 2023). This would allow teaching quality measures to encompass individual and unique interactions with students, thus enhancing their relevance at the classroom level. Moreover, future methodological studies could explore the role of teaching quality and learning processes, particularly by using qualitative interviews, to develop more adequate measures.

4.4 Implications for teaching practice

The primary focus of our study was to improve the conceptual understanding of teaching and its effects on student outcomes for future research on teaching quality. The findings demonstrate that we are a long way from fully understanding the mediating mechanisms that underlie how teaching affects student outcomes. While these factors make it more difficult to suggest implications for practice than, for example, with an intervention study, we do believe that the results are relevant to teaching practice in two ways.

First, that the study found no mediation effects at the classroom level, but several at the student level, suggests that teachers might need shift their focus from the class to the student. This requires a more adaptive approach to teaching, one that is responsive to the evolving dynamics of the class and addresses not just the collective needs of the class but also the unique needs of each student (Vieluf & Klieme, 2023). While this is not an original recommendation – researchers have been discussing the idea, mostly at a theoretical level, for decades – our study provides supporting empirical evidence. Clearly, this is a challenging remit for teachers. Support could include developing formative tools to help teachers gather and interpret student perceptions of teaching and use of learning materials, and providing concrete guidance on how to incorporate this information into daily lesson plans (Decristan et al., 2015; Pinger et al., 2018).

Second, results suggest that the mechanisms through which teaching shapes learning are far more complex than teaching effectiveness researchers had hitherto hypothesized. Not only has our study uncovered a more intricate array of mediation pathways within the original model than previously identified, but it also suggests that an expansion to encompass adjacent theoretical frameworks such as EVT (Wigfield & Eccles, 1992) may reveal yet more mediators. This complexity means that there can be no standard teaching "recipes" that work for all students (see Vieluf, 2022). Of course teachers, especially trainees, find recipes appealing but these results suggest that teaching is too complex and constrained by context for such prescriptions.

To conclude, our study highlights the importance of focusing on the individual student's use of learning opportunities, resonating with constructivist principles that emphasize the importance of the individual's construction of knowledge (Aebli, 2011; Piaget, 1992). This in turn underscores the value of an adaptive and flexible approach to teaching; one that is responsive to the evolving dynamics of the classroom (Vieluf & Klieme, 2023).

4.5 Limitations and Future Directions

First, contrary to our expectations, cognitive activation was not related to time-on-task at either the classroom or student level. Looking at the operationalization of the constructs in more detail, it becomes evident that our study operationalized cognitive activation specifically as teachers providing tasks that require critical thinking and presenting problems with no obvious solutions, whereas time-on-task was defined and operationalized as paying attention during the mathematics lesson, but not specifically when solving complex tasks or undertaking critical thinking. When comparing the operationalization of the constructs, it appears that time-on-task was more generally operationalized than cognitive activation. In a similar vein, the variables at T2 referred to specific mathematics lessons on quadratic equations, whereas teaching quality and the mediators referred to mathematics in general. Those issues with operationalization could have resulted in greater variability in the way students interpret and respond to the items, which may not have been evident when analyzing the results. Some students might usually listen to instructions and pay attention during the lessons but perhaps not be particularly attentive when solving complex problems or vice versa. Some students may also be more interested or more successful in some academic domains than in others (Jansen et al., 2019). Considering all these issues, it would be fruitful to compare the effects by using different operationalizations of the constructs, ideally within one study.

Second, in the TALIS Video Study space restrictions in the questionnaire meant the item for assessing some variables did not permit a detailed investigation of the subdimensions. For example, need satisfaction was assessed by three items, one item for each need: autonomy, competence, and relatedness. Research in SDT has begun to focus on the negative impact of need frustration, not just the positive effect of need satisfaction. We suggest that future studies on the TBD model incorporate these theoretical developments by using more comprehensive and well-established questionnaires such as the

Basic Psychological Needs Satisfaction Scales (BPNSS; Deci & Ryan, 2000; Gagné, 2003), the Balanced Measure of Psychological Needs (BPMN; Sheldon & Hilpert, 2012), and the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS; Chen et al., 2015, Van der Kaap-Deeder et al., 2020).

Third, the achievement test used in the TALIS Video Study focused on low- to medium-level cognitive demands such as memorization, procedures and simple applications, and did not adequately assess students' high-level thinking such as using multiple representations and modelling authentic situations. This limitation was due in part to the difficulty of optimizing the achievement test for the diverse curricula in the countries participating in the TALIS Video Study (Herbert et al., 2022). It is also important to consider that the way the test was administered during the study differed from usual classroom procedures and this may have affected the performance of some students. For example, some students may have been less attentive or more anxious during the test, which could have influenced their performance. It is therefore important to carefully consider the selection and adaptation of measures to accurately capture the constructs of interest in research and to also consider the potential impact of situational factors on performance.

Fourth, the sample in our study was recruited from secondary schools in Germany and is highly selective as teachers decided in whether to participate in the study or not. Therefore, the results cannot be generalized to the secondary schools in Germany in general, and, even more so, not to other age groups (e.g., primary school or university) or other countries. The results also differed when the effects of the three basic dimensions of teaching quality on student outcomes was analyzed for the different countries which participated in the TALIS Video Study (Herbert et al., 2022). Therefore, cross-cultural studies should investigate mediating effects to strengthen the generalizability of our findings.

The fifth limitation relates to measurement perspective, which can have an impact on study results (e.g., Zee et al., 2013). This study used student ratings, which are considered valid and are commonly used in the field (Appleton et al., 2008; De Jong & Westerhof, 2001; Fredricks, 2022; Lüdtke et al., 2009). Student ratings of teaching quality have been found in some studies to be a better predictor of student variables than teacher and observer ratings (e.g., Kunter & Baumert, 2006; Styck et al., 2020; Wagner et al., 2016). However, observer ratings have the potential to be a more objective measure of teaching quality (Clausen, 2002) and it may be that using student data for assessing teaching quality as well as student learning processes introduces a risk of common method bias, particularly in correlational analyses (Podsakoff et al., 2003; 2012). Although it is difficult to identify this bias empirically, future studies might consider using a marker variable which is theoretically unrelated to the other variables of the study (Williams et al., 2010). Self-report surveys might also be affected by what is socially desirable, leading to over- or under-representation of the participant's actual behavior (Fredricks, 2022). We suggest future studies incorporate multiple perspectives to measure variables and, for simplicity, when comparing mediating effects according to rater perspectives, focus only on any specific paths within the same study (Fauth et al., 2020).

To summarize, neither the direct nor the indirect effect models in our study provide clear answers about the hypothesized relationships. There could be multiple reasons for these findings. Our results reveal the critical importance of certain choices made while designing and analyzing a study. It seems that the conceptual sequence of the variables, the choice of correlational vs. longitudinal evidence, and the level of analysis all have an impact on the results. Our finding is in line with recent reviews that have also revealed inconsistent results (see Alp Christ et al., 2022; Praetorius et al., 2018). Thus, one important take home message is that current quantitative results on the direct and indirect effects of teaching quality on student outcomes are not easy to interpret. Instead of considering the entire chain at once, it might be more productive to focus exclusively on understanding the interplay between teaching quality and student learning processes better for a while (see also Hiebert & Stigler, 2023).

5. Conclusions

This study is the first to investigate relationships within the entire TBD model using a longitudinal design. It does so by enriching the TBD model with well-established cognitive and motivational theories. The multilevel mediation analyses using both correlational and longitudinal designs revealed varied results which depended on study design and level of analysis and once again highlighted the complexity of the relationships between teaching quality, student learning processes, and student outcomes (see also Alp Christ et al., 2022). Our study contributes to the literature by supporting some of the assumptions of the TBD model and finding new paths between teaching quality and student outcomes. In line with recent appeals in the field (Praetorius & Charalambous, 2023; Vieluf & Klieme, 2023), our study advocates for augmenting the current model with supplementary theories pertaining to cognition, motivation, and effort, to advance the field.

Keypoints

- The assumptions of the TBD model are revisited and expanded using leading motivational and cognitive theories.
- First longitudinal investigation of the entire TBD model, integrating new possible mediating paths.
- Multilevel mediation analyses show diverse findings for direct and indirect effects, highlighting model intricacies.
- Conceptual and methodological choices can have a significant influence on the results.

Supplementary Material

Supplementary material for this article can be found online.

Funding

Ayşenur Alp Christ is funded by the Swiss Government Excellence Scholarship (ESKAS No. 2019.0503). The TALIS Video Study Germany was supported by the Leibniz Association.

Author Contributions

Ayşenur Alp Christ: Conceptualization, literature review, writing – original draft, writing – review & editing, data analysis, visualization, supplementary materials. **Vanda Capon-Sieber:** Conceptualization, literature review, writing – original draft, supervision, writing – review & editing. **Carmen Köhler:** Supervision of the data analyses, writing – review & editing. **Eckhard Klieme:** Design of the TALIS-Video study, data collection and curation, writing – review & editing. **Anna-Katharina Praetorius:** Conceptualization, design of the study, data collection and curation, supervision, writing – review & editing.

Appendix

Items assessed in the TALIS Video Study Student Questionnaire

Cognitive activation with discourse (T1) (1= never or almost never , 2 = occasionally, 3 = frequently, 4 = always)

Our mathematics teacher presents tasks for which there is no obvious solution.
Our mathematics teacher presents tasks that require us to apply what we have learned to new contexts.
Our mathematics teacher gives tasks that require us to think critically.
Our mathematics teacher asks us to decide on our own procedures for solving complex tasks.
Our mathematics teacher gives us opportunities to explain our ideas.
Our mathematics teacher encourages us to question and critique arguments made by other students.
Our mathematics teacher requires us to engage in discussions among ourselves.

Classroom management (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)

When the lesson begins, our mathematics teacher has to wait quite a long time for us to quieten down.
We lose quite a lot of time because of students interrupting the lesson.
There is much disruptive noise in this classroom.
In our teacher's class, we are aware of what is allowed and what is not allowed.
In our teacher's class, we know why certain rules are important.
Our teacher manages to stop disruptions quickly.
Our teacher reacts to disruptions in such a way that the students stop disturbing learning.
In our teacher's class, transitions from one phase of the lesson to the other (e.g., from <class> discussions to individual work) take a lot of time.

Student Support (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)

Our mathematics teacher gives extra help when we need it.
Our mathematics teacher continues teaching until we understand.
Our mathematics teacher helps us with our learning.
Our mathematics teacher makes me feel confident in my ability to do well in the <course>.
Our mathematics teacher listens to my view on how to do things.
I feel that our mathematics teacher understands me.
Our mathematics teacher makes me feel confident in my ability to learn the material.
Our mathematics teacher provides me with different alternatives (e.g. learning materials or tasks).
Our mathematics teacher encourages me to find the best way to proceed by myself.
Our mathematics teacher lets me work on my own.
Our mathematics teacher appreciates it when different solutions come up for discussion.

Depth of processing (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)

I keep thinking about tasks until I really understand them.
I think intensively about the mathematical content.
I develop my own ideas regarding the topic taught.

Need satisfaction (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)

I feel that I can decide on things on my own.
I feel understood by my mathematics teacher.
I feel confident in my ability to learn this material.

Time-on-task (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)

I pay attention in mathematics class.
I listen to the instruction given in class.
I let my mind wander during the lessons.

Interest (T1) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)
I am interested in mathematics.
I often think that what we are talking about in my mathematics class is interesting.
After mathematics class I am often already curious about the next mathematics class.

Interest (T2) (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree)
I was interested in the topic of quadratic equations.
I often thought that what we were talking about in my mathematics class during the unit on quadratic equations was interesting.
After my mathematics class on the topic of quadratic equations I was often already curious about the next mathematics class.

References

Adachi, P., & Willoughby, T. (2015). Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: Implications for psychological science. *European Journal of Developmental Psychology, 12,* 116–128. https://doi.org/10.1080/17405629.2014.963549

Aebli, H. (2011). *Zwölf Grundformen des Lehrens: Eine allgemeine Didaktik aufpsychologischer Grundlage. Medien und Inhalte didaktischer Kommunikation, der Lernzyklus* [Twelve basic forms of teaching: A general didactics based on psychology: Media and content of didactic communication, the learning cycle] (14. Auflage). Klett-Cotta.

Ahmadi, A., Noetel, M., Parker, P., Ryan, R.M., Ntoumanis, N., Reeve, J., Beauchamp, M.,Dicke, T., Yeung, A., Ahmadi, M., Bartholomew, K., Chiu, T.K.F., Curran, T., Erturan, G., Flunger, B., Frederick, C., Froiland, J. M., González-Cutre, D., Haerens, L., . . . Lonsdale, C. (2023). A classification system for teachers' motivational behaviors recommended in self-determination theory interventions. *Journal of Educational Psychology*. *115*(8), 1158–1176. https://doi.org/10.1037/edu0000783

Ahn, I., Ming Chiu, M., & Patrick, H. (2021). Connecting teacher and student motivation: Student-perceived teacher need-supportive practices and student need satisfaction. *Contemporary Educational Psychology*, *64,* 101950. https://doi.org/10.1016/j.cedpsych.2021.101950

Alp Christ, A., Capon-Sieber, V., Grob, U.W., & Praetorius, A. K. (2022). Learning processes and their mediating role between teaching quality and student achievement: A systematic review. *Studies in Educational Evaluation*, *75*, 101209. https://doi.org/10.1016/j.stueduc.2022.101209

Appleton, J.J., Christenson, S.L., & Furlong, M.J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45,* 369–386. https://doi.org/10.1002/pits.20303

Appleton, J. J., Christenson, S.L., Kim, D., & Reschly, A.L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*(5), 427–445. https://doi.org/10.1016/j.jsp.2006.04.002

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64, Part 1*(6), 359–372. https://doi.org/10.1037/h0043445

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K.W. Spence & J.T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory.* (Vol. 2, pp. 89-195). Academic Press.

Badri, R., Amani-Saribaglou, J., Ahrari, G., Jahadi, N., & Mahmoudi, H. (2014). School culture, basic psychological needs, intrinsic motivation and academic achievement: Testing a casual model. *Mathematics Education Trends and Research, 4,* 1–13. https://doi.org/10.5899/2014/metr-00050

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal, 47*(1), 133–180. https://doi.org/10.3102/0002831209345157

Begrich, L., Praetorius, A. K., Decristan, J., Fauth, B., Göllner, R., Herrmann, C., Kleinknecht, M., Taut, S., & Kunter, M. (2023). Was tun? Perspektiven für eine Unterrichtsqualitätsforschungder Zukunft. [What to do? Perspectives for a teaching quality research of the future]. *Unterrichtswissenschaft*, *51*, 63–97. https://doi.org/10.1007/s42010-023-00163-4

Berliner, D. C. (2002). Comment: Educational research: The hardest science of all. *Educational Researcher*, *31*(8), 18–20. https://doi.org/10.3102/0013189X031008018

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education, 84*(6), 740–756. https://doi.org/10.1002/1098237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*, 111–127. https://doi.org/10.1037//0003-066X.57.2.111
Boggiano, A. K., Main, D. S., & Katz, P. A. (1988). Children's preference for challenge: The role of perceived competence and control. *Journal of Personality and Social Psychology*, *54*(1), 134–141. https://psycnet.apa.org/doi/10.1037/0022-3514.54.1.134
Boggiano, A. K., Flink, C., Shields, A., Seelbach, A., & Barrett, M. (1993). Use of techniques promoting students' self-determination: Effects on students' analytic problem-solving skills. *Motivation and Emotion, 17*(4), 319–336. https://doi.org/10.1007/BF00992323
Boston, M. D., & Candela, A. G. (2018). The Instructional Quality Assessment as a tool for reflecting on instructional practice. *ZDM, 50*(3), 427–444. https://doi.org/10.1007/s11858-018-0916-6
Böheim, R., Knogler, M., Kosel, C., & Seidel, T. (2020). Exploring student hand-raising across two school subjects using mixed methods: An investigation of an everyday classroom behavior from a motivational perspective. *Learning and Instruction, 65,* 101250. https://doi.org/10.1016/j.learninstruc.2019.101250
Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*, 109–131. https://doi.org/10.1146/annurev.ps.40.020189.000545
Brophy, J. (2000). *Teaching. Educational practices series: Vol. 1.* InternationalAcademy of Education.
Brophy, J. (2006). History of research on classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 17–43). Lawrence Erlbaum.
Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. https://doi.org/10.1037/0022-3514.84.4.822
Burić, I., & Kim, L. E. (2020). Teacher self-efficacy, instructional quality, and student motivational beliefs: An analysis using multilevel structural equation modeling. *Learning and Instruction*, *66*, 101302. https://doi.org/10.1016/j.learninstruc.2019.101302
Butler, R., & Shibaz, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction, 18*(5), 453–467. https://doi.org/10.1016/j.learninstruc.2008.06.004
Chan, S., & Yuen, M. (2014). Personal and environmental factors affecting teachers' creativity-fostering practices in Hong Kong. *Thinking Skills and Creativity, 12*, 69-77. https://doi.org/10.1016/j.tsc.2014.02.003
Charalambous, C. Y., & Praetorius, A.-K. (2020). Creating a forum for researching teaching and its quality more synergistically. *Studies in Educational Evaluation*, *67*, 100894. https://doi.org/10.1016/j.stueduc.2020.100894
Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., & Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*(2), 216–236. https://doi.org/10.1007/s11031-014-9450-1
Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables. *Organizational Research Methods, 11*(2), 296–325. https://doi.org/10.1177/1094428107300343
Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243.https://doi.org/10.1080/00461520.2014.965823
Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperativelearning: How to improve student problem solving and time on-task. *American Educational Research Journal, 41*(2), 365–399. https://doi.org/10.3102/00028312041002365
Clausen, M. (2002). Unterrichtsqualität: Eine Frage der Perspektive? [Instructional quality: A question of perspectives?]. Waxmann.

Clifford, M. (1990). Students need challenge, not easy success: Only by teaching students to tolerate failure for sake of true success can educators control the national epidemic of "educational suicide". *Educational Leadership, 48*, 22–26.
Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist, 30*, 116–127. https://doi.org/10.1037/h0076829
De Corte, E. (1995). Fostering cognitive growth: A perspective from research on Mathematics learning and instruction. *Educational Psychologist, 30*(1), 37–46. https://doi.org/10.1207/s15326985ep3001_4
De Corte, E. (2004). Mainstreams and perspectives in research on learning (mathematics) from instruction. *Applied Psychology, 53*(2), 279–310. https://doi.org/10.1111/j.1464-0597.2004.00172.x
De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research*, *4*(1), 51–85. https://doi.org/10.1023/A:1011402608575
Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality, 62*(1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x
Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11,* 227–268. https://doi.org/10.1207/S15327965PLI1104_01
Deci, E. L., & Ryan, R. M. (Eds.). (2002). Handbook of self-determination research. The University of Rochester Press.
Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., Hondrich, A.L., Rieser, S., Hertel, S. & Hardy, I. (2015). Embedded formative assessment and classroom process quality: How do they interact in promoting science understanding? *American Educational Research Journal*, *52*(6), 1133-1159. https://doi.org/10.3102/0002831215596412
Diederich, J., & Tenorth, H.-E. (1997). *Theorie der Schule: Ein Studienbuch zu Geschichte, Funktionen und Gestaltung* [Theory of school: A study book on history, functions and design]. Berlin, Germany: Cornelsen Scriptor. https://doi.org/10.25656/01:11675
Dorfner, T., Förtsch, C., & Neuhaus, B. J. (2018). Effects of three basic dimensions of instructional quality on students' situational interest in sixth-grade biology instruction. *Learning and Instruction, 56*, 42–53. https://doi.org/10.1016/j.learninstruc.2018.03.001
Doyle, W. (1983). Academic work. *Review of Educational Research, 53*, 159–200. https://doi.org/10.3102/00346543053002159
Doyle, W. (1986). Classroom organization and management. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 392–431). MacMillan.
Driscoll, M. P. (2005). *Psychology of learning for instruction.* Peterson.
Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist, 36*(2), 103–112. https://doi.org/10.1207/S15326985EP3602_5
Evertson, C. M. (1989). Improving elementary classroom management: A school-based training program for beginning the year. *The Journal of Educational Research, 83*(2), 82–90. https://doi.org/10.1080/00220671.1989.10885935
Evertson, C. M., & Harris, A. M. (1992). What we know about managing classrooms. *Educational Leadership, 49,* 74–78.
Fauth, B., Decristan, J., Decker, A.-T., Büttner, G., Hardy, I., Klieme, E., & Kunter, M. (2019). The effects of teacher competence on student outcomes in elementary science education: The mediating role of teaching quality. *Teaching and Teacher Education, 86*, 102882. https://doi.org/10.1016/j.tate.2019.102882
Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29,* 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Fauth, B., Göllner, R., Lenske, L., Praetorius, A., & Wagner, W. (2020). Who sees what? Theoretical considerations on the measurement of teaching quality from different perspectives. *Zeitschrift für Pädagogik, 66*(1), 138–155. https://doi.org/10.25656/01:25870

Finn, J. D., & Zimmer, K. S. (Eds.). (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.) *Handbook of research on student engagement* (pp 97–131). Springer. https://doi.org/10.1007/978-1-4614-2018-7_5

Fisher, C., Berliner, D., Filby, N., Marliave, R., Cahen, L., & Dishaw, M. (1981). Teaching behaviors, academic learning time, and student achievement: An overview. *The Journal of Classroom Interaction, 17*(1), 2–15.

Förtsch, C., Werner, S., Dorfner, T., Kotzebue, L. von, & Neuhaus, B. J. (2017). Effects of cognitive activation in biology lessons on students' situational interest and achievement. *Research in Science Education, 47*(3), 559–578. https://doi.org/10.1007/s11165-016-9517-y

Förtsch, C., Werner, S., Kotzebue, L. von, & Neuhaus, B. J. (2018). Effects of high-complexity and high-cognitive-level instructional tasks in biology lessons on students' factual and conceptual knowledge. *Research in Science & Technological Education, 36*(3), 353–374. https://doi.org/10.1080/ 02635143.2017.1394286

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 763–782). Springer US. https://doi.org/10.1007/978-1-4614-2018-7_37

Fredricks, J. A. (2022). The measurement of student engagement: Methodological advances and comparison of new self-report instruments. In A. Reschly & S. Christenson (Eds.), *Handbook of research on student engagement* (pp. 597–616). Springer International Publishing. https://doi.org/10.1007/ 978-3-031-07853-8

Gagné, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion, 27*, 199–223. https://doi.org/10.1023/A:1025007614869

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods, 19*(1), 72–91. https://doi.org/10.1037/a0032138

Good, T. L., & Brophy, J. E. (2003). *Looking in classrooms* (9th ed.). Pearson Education.

Grabinger, R. S., & Dunlap, J. C. (1995). Rich environments for active learning: A definition. *Research in Learning Technology* , *3*(2), 5–34. https://doi.org/10.3402/rlt.v3i2.9606

Grabinger, S., Dunlap, J. C., & Duffield, J. A. (1997). Rich environments for active learning in action: Problem-based learning. *Research in Learning Technology*, *5*(2), 5–17. https://doi.org/10.3402/rlt.v5i2.10558

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie Canadienne*, *49*(3), 233–240. https://doi.org/10.1037/a0012758

Hachfeld, A., & Lazarides, R. (2020). The relation between teacher self-reported individualization and student-perceived teaching quality in linguistically heterogeneous classes: An exploratory study. *European Journal of Psychology of Education*, *36*, 1159–1179. https://doi.org/10.1007/s10212-020-00501-5

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

Heckhausen, H. (1991). *Motivation and Action* (P. K. Leppman, Trans). Springer-Verlag.

Helm, C. (2016). Zentrale Qualitätsdimensionen von Unterricht und ihre Effekte auf Schüleroutcomes im Fach Rechnungswesen. [Key quality dimensions of teaching and their effects on student outcomes in accounting] *Zeitschrift Für Bildungsforschung, 6*(2), 101–119. https://doi.org/10.1007/s35834-016-0154-3

Helmke, A. (2012). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* [Teaching quality and teacher professionalism: Diagnosis, evaluation, and improvement of teaching] (4. überarbeitete Aufl.). Klett/Kallmeyer. https://books.google.ch/books?id=MHdxOwAACAAJ

Herbert, B., Fischer, J., & Klieme, E. (2022). How valid are student perceptions of teaching quality across education systems? *Learning and Instruction*, *82*, 101652. https://doi.org/10.1016/j.learninstruc .2022.101652
Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 371–404). Information Age Pub.
Hiebert, J., & Stigler, J. W. (2023). Creating practical theories of teaching. In A.-K. Praetorius & C. Y. Charalambous (Eds.), *Theorizing teaching: Current status and open issues* (pp. 23–56). Springer.
Hill, P., & Rowe, K. (1996). Multilevel modeling in school effectiveness research. *School Effectiveness and School Improvement, 7*, 1–34. https://www.tandfonline.com/doi/pdf/10.1080/0924345960070101
Hospel, V., & Galand, B. (2016). Are both classroom autonomy support and structure equally important for students' engagement? A multilevel analysis. *Learning and Instruction, 41*, 1–10. https://doi.org/10.1016/j.learninstruc.2015.09.001
Hox, J., & McNeish, D. (2020). Small samples in multilevel modeling. In R. Van de Schoot, & M. Miočević (Eds.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 215–225). Routledge. https://doi.org/10.4324/9780429273872-18 .
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1– 55. https://doi.org/10.1080/10705519909540118
Hughes, J. N., Luo, W., Kwok, O. M., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology*, *100*(1), 1–14. https://doi.org/10.1037/0022-0663.100.1.1
Illies, J. J., & Reiter-palmon, R. (2004). The effects of type and level of personal involvement on information search and problem solving. *Journal of Applied Social Psychology*, *34*(8), 1709–1729. https://doi.org/10.1111/j.1559-1816.2004.tb02794.x
Jang, H., Kim, E. J., & Reeve, J. (2012). Longitudinal test of self-determination theory's motivation mediation model in a naturally occurring classroom context. *Journal of Educational Psychology*, *104*(4), 1175–1188. https://doi.org/10.1037/a0028089
Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2019). The dimensional structure of students' self-concept and interest in science depends on course composition. *Learning and Instruction*, *60*, 20–28. https://doi.org/10.1016/j.learninstruc.2018.11.001
Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 8*(3), 325–352. https://doi.org/10.1207/S15328007SEM0803_1
Kiemer, K., Gröschner, A., Kunter, M., & Seidel, T. (2018). Instructional and motivational classroom discourse and their relationship with teacher autonomy and competence support—findings from teacher professional development. *European Journal of Psychology of Education, 33*(2), 377–402. https://doi.org/10.1007/s10212-016-0324-7
Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht: Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras".[Quality dimensions and effectiveness of mathematics education: theoretical foundations and selected results of the "Pythagoras" project.] In M. Prenzel & L. Allolio-Näcke (Eds.), *Untersuchungen zur Bildungsqualität von Schule: Abschlussbericht des DFG-Schwerpunktprogramms* (pp. 127–146). Waxmann.
Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janík & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Waxmann.

Klieme, E., & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht [Instructional quality from a student perspective: Culture-specific profiles, regional differences, and associations with effects of instruction]. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann, & M. Weiß (Eds.), *PISA 2000 — Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (pp. 333–359). Leske + Budrich; VS Verlag für Sozialwissenschaften.

Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik: Outcome-orientierte Messung und Prozessqualität des Unterrichts. [Empirical classroom research and subject didactics: Outcome-oriented measurement and process quality of teaching. *Zeitschrift Für Pädagogik, 54*(2), 222–237. https://doi.org/10.25656/01:4348

Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: «Aufgabenkultur» und Unterrichtsgestaltung [Mathematics teaching in lower secondary schools: "Task culture" and lesson design]. In E. Klieme, & J. Baumert (Eds.), *TIMSS – Impulse für Schule und Unterricht* (pp. 43-57). Bundesministerium für Bildung und Forschung.

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press

Kounin, J. S. (1970a). *Discipline and group management in classrooms*. Holt Rinehart &Winston.

Kounin, J. S. (1970b). Observing and delineating technique of managing behavior in classrooms. *Journal of Research and Development in Education*, *4*(1), 62–67.

Kuger, S. (2016). Curriculum and learning time in international school achievement studies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Assessing contexts of learning* (pp. 395–422). Springer International Publishing. https://doi.org/10.1007/978-3-319-45357-6_16

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environment Research 9*, 231–251. https://doi.org/10.1007/s10984-006-9015-7

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction, 17*(5), 494–509. https://doi.org/10.1016/j.learninstruc.2007.09.002

Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts [Psychology of Teaching]. StandardWissen Lehramt: Vol. 3895.* Ferdinand Schöningh. https://doi.org/10.36198/9783838538952

Lazarides, R., & Buchholz, J. (2019). Student-perceived teaching quality: How is it related to different achievement emotions in mathematics classrooms? *Learning and Instruction, 61*, 45–59. https://doi.org/10.1016/j.learninstruc.2019.01.001

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642

León, J., Medina-Garrido, E., & Núñez, J. L. (2017). Teaching quality in math class: The development of a scale and the analysis of its relationship with engagement and achievement. *Frontiers in Psychology*, *8*, 895. https://doi.org/10.3389/fpsyg.2017.00895

Li, H., Liu, J., Zhang, D., & Liu, H. (2020). Examining the relationships between cognitive activation, self-efficacy, socioeconomic status, and achievement in mathematics: A multi-level analysis. *The British Journal of Educational Psychology, 91,* 101–126. https://doi.org/10.1111/bjep.12351

Lietaert, S., Roorda, D., Laevers, F., Verschueren, K., & De Fraine, B. (2015). The gender gap in student engagement: The role of teachers' autonomy support, structure, and involvement. *British Journal of Educational Psychology, 85*(4), 498–518. https://doi.org/10.1111/bjep.12095

Lipowsky, F., & Hess, M. (2019). Warum es manchmal hilfreich sein kann, das Lernen schwerer zu machen – Kognitive Aktivierung und die Kraft des Vergleichens [Why it can sometimes be helpful to make learning harder – Cognitive activation and the power of comparison]. In K. Schöppe & F. Schulz (Hrsg.), *Kreativität & Bildung –Nachhaltiges Lernen* (s. 77–132). kopaed.

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the

Pythagorean Theorem. *Learning and Instruction, 19*(6), 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001
Lotz, M. (2016). Grundlagen des Unterrichtens und der Unterrichtsforschung [Fundamentals of teaching and classroom research]. In M. Lotz (Ed.), *Kognitive Aktivierung im Leseunterricht der Grundschule: Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr* (pp. 7–22). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-10436-8_2
Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modelling. *Contemporary Educational Psychology, 34*(2), 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001
Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. Methodology: European *Journal of Research Methods for the Behavioral and Social Sciences, 1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86
Mackinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research, 39*(1), 99–128. https://doi.org/10.1207/s15327906mbr3901_4
Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., et al. (2012). Classroom climate and contextual effects: conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*, 106–124. http://dx.doi.org/10.1080/00461520.2012.670488
McCaffrey, D. F., Castellano K. E., van Essen, T. (2020). Student test development. In Global Teaching InSights: Technical report. Section II: Instrument development. OECD.
McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.
Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85,* 424–436. https://doi.org/10.1037/0022-0663.85.3.424
Mouratidis, A., Barkoukis, V., & Tsorbatzoudis, C. (2015). The relation between balanced need satisfaction and adolescents' motivation in physical education. *European Physical Education Review, 21*(4), 421–431. https://doi.org/10.1177/1356336X15577222
Mouratidis, A., Vansteenkiste, M., Michou, A., & Lens, W. (2013). Perceived structure and achievement goals as predictors of students' self-regulated learning and affect and the mediating role of competence need satisfaction. *Learning and Individual Differences, 23*, 179–186. https://doi.org/10.1016/j.lindif.2012.09.001
Murray, C., & Pianta, R. C. (2007). The importance of teacher-student relationships for adolescents with high incidence disabilities. *Theory Into Practice*, *46*(2), 105–112. https://doi.org/10.1080/00405840701232943
OECD. (2020). *Global teaching insights: A video study of teaching*. OECD. https://doi.org/ 10.1787/20d6f36b-en
Pauli, C., & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. [From international comparative video surveys to video-based teaching research and development] *Zeitschrift für Pädagogik, 52,* 774−798. https://doi.org/10.25656/01:4488
Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341. https://doi.org/10.1007/s10648-006-9029-9
Petty, R. E., & Cacioppo, J. T. (1986): The elaboration likelihood model of persuasion. In L. Berkowitz (Eds.), 19, *Advances in experimental social psychology* (pp. 123 – 205). New York: Academic Press. https://doi.org/10.1016/S0065-2601(08)60214-2
Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135–146. https://doi.org/10.1086/208954
Piaget, J. (1992). *Psychologie der Intelligenz [Psychology of the Intelligence]*. Klett-Cotta.

Pianta, R. C., & Steinberg, M. (1992). Teacher–child relationships and the process of adjusting to school. *New Directions for Child and Adolescent Development, 57,* 61–80. https://doi.org/10.1002/cd.23219925706
Pinger, P., Rakoczy, K., Besser, M., & Klieme, E. (2018). Interplay of formative assessment and instructional quality—interactive effects on students' mathematics achievement. *Learning Environments Research*, *21*, 61-79. https://doi.org/10.25656/01:17407
Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879
Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*, 539–569. https://doi.org/10.1146/annurev-psych-120710-100452
Praetorius, A.-K., & Charalambous, C. Y. (2023). Where are we on theorizing teaching? A literature overview. In A.-K. Praetorius & C. Y. Charalambous (Eds.), *Theorizing teaching: Current status and open issues* (pp. 1–22). Springer. https://doi.org/10.1007/978-3-031-25613-4_1
Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. *ZDM, 50*(3), 407–426. https://doi.org/10.1007/s11858-018-0918-4
Praetorius, A.-K., Fischer, J., & Klieme, E. (2020b). Teacher and student questionnaire development. In Global Teaching InSights Technical report. Section II: Instrument development. Paris: OECD Publishing.
Praetorius, A. K., Grünkorn, J., & Klieme, E. (2020a). Towards developing a theory of generic teaching quality: origin, current status, and necessary next steps regarding the three basic dimensions model. *Zeitschrift für Pädagogik. Beiheft*, *66*(1), 15–36. https://doi.org/10.3262/ZPB2001015
Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K., & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction, 31*, 2–12. https://doi.org/10.1016/j.learninstruc.2013.12.002
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.
https://doi.org/10.3758/BRM.40.3.879
Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2011). Alternative methods for assessingmediation in multilevel data: The advantages of multilevel SEM. *Structural Equation Modeling, 18*, 161–182. https://doi.org/10.1080/10705511.2011.557329
R Development Core Team. (2020). *R: A language and environment for statistical computing* [Computer software]. http://www.R-project.org/
Rakoczy, K. (2006). Motivationsunterstützung im Mathematikunterricht. Zur Bedeutung von Unterrichtsmerkmalen für die Wahrnehmung von Schülerinnen und Schüler. [Motivational support in mathematics education. On the importance of instructional features for the perception of students]. *Zeitschrift Für Pädagogik, 52*(6), 822–843. https://doi.org/10.25656/01:4490
Rakoczy, K., & Pauli, C. (2006). Hoch inferentes Rating. Beurteilung der Qualitätunterrichtlicher Prozesse [High-inference rating. Assessment of instructional quality]. In I. Hugener, C. Pauli, & K. Reusser (Eds.), *Video-analysen. Dokumentation der Erhebungs- und Auswertungsinstrumente zurschweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhaltenund mathematisches Verständnis". Materialien zur Bildungsforschung*. GFPF, 15, 206–233.
Reeve, J. (2006). Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit. *The Elementary School Journal*, *106*(3), 225–236. http://dx.doi.org/10.1086/501484
Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. https://doi.org/10.1037/0022-0663.98.1.209

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159–175. https://doi.org/10.1080/00461520903028990
Reeve, J. (2016). Autonomy-supportive teaching: What it is, how to do it. In W. C. Liu, J. C. K. Wang, and R. M. Ryan (Eds.), *Building autonomous learners: Perspectives from research and practice using self-determination theory* (pp. 129–152). Singapore: Springer Singapore.
Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195). Academic Press. https://doi.org/10.1016/B978-012750053-9/50009-7
Reusser, K., Pauli, C., & Waldis, M. (Eds.). (2010). *Unterrichtsgestaltung und Unterrichtsqualität: Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht*. *[Lesson design and Teaching quality: Results of an international and Swiss video study on mathematics teaching]*. Waxmann.
Richter, M., Gendolla, G., & Wright, R. A. (2016). Three decades of research on motivational intensity theory. In *Advances in Motivation Science* (Vol. 3, pp. 149–186). https://doi.org/10.1016/bs.adms.2016.02.001
Rieser, S., Naumann, A., Decristan, J., Fauth, B., Klieme, E., & Büttner, G. (2016). The connection between teaching and learning: Linking teaching quality and metacognitive strategy use in primary school. *British Journal of Educational Psychology*, *86*(4), 526–545. https://doi.org/10.1111/bjep.12121
Rosseel, Y. (2012). lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, *48* (2), 1–36. https://doi.org/10.18637/jss.v048.i02
Ruiz-Alfonso, Z., & León, J. (2017). Passion for math: Relationships between teachers' emphasis on class contents usefulness, motivation and grades. *Contemporary Educational Psychology*, *51*, 284–292. https://doi.org/10.1016/j.cedpsych.2017.08.010
Ruiz-Alfonso, Z., & León, J. (2019). Teaching quality: relationships between passion, deep strategy to learn, and epistemic curiosity. *School Effectiveness and School Improvement*, *30*(2), 212–230. https://doi.org/10.1080/09243453.2018.1562944
Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/ 10.1037//0003-066X.55.1.68
Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). University of Rochester Press.
Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wentzel & A. Wigfield (Eds.), *Educational psychology handbook series. Handbook of motivation at school* (pp. 171–196). Routledge.
Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* Guilford Publications.
Ryan, R. M., & Powelson, C. L. (1991). Autonomy and relatedness as fundamental to motivation and education. *The Journal of Experimental Education, 60*(1), 49–66. https://doi.org/ 10.1080/00220973.1991.10806579
Savalei, V., & Rosseel, Y. (2021). Computational options for standard errors and test statistics with incomplete normal and nonnormal data in SEM. *Structural Equation Modeling: A Multidisciplinary Journal, 29*(2), 163–181. https://doi.org/10.31234/osf.io/wmuqj
Schlesinger, L., & Jentsch, A. (2016). Theoretical and methodological challenges in measuring instructional quality in mathematics education using classroom observations. *ZDM*, *48*, 29–40. https://doi.org/ 10.1007/s11858-016-0765-0
Schlesinger, L., Jentsch, A., Kaiser, G., König, J., & Blömeke, S. (2018). Subject-specific characteristics of instructional quality in mathematics education. *ZDM*, *50*, 475–490. https://doi.org/10.1007/s11858-018-0917-5
Schukajlow, S., Achmetli, K., & Rakoczy, K. (2019). Does constructing multiple solutions for real-world problems affect self-efficacy? *Educational Studies in Mathematics, 100*(1), 43–60.

https://doi.org/10.1007/s10649-018-9847-y
Schukajlow, S., & Krug, A. (2014). Do multiple solutions matter? Prompting multiple solutions, interest, competence, and autonomy. *Journal for Research in Mathematics Education, 45*, 497–533. https://doi.org/10.5951/jresematheduc.45.4.0497
Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM–Mathematics Education, 55*, 249–267. https://doi.org/10.1007/s11858-022-01463-2
Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie: Integration von Struktur- und Prozessparadigma [The opportunity-use model in instructional psychology: Integrating structure and process paradigms]. *Zeitschrift für Pädagogik*, *60*(6), 850–866. https://doi.org/10.25656/01:14686
Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499. https://doi.org/10.3102/0034654307310317
Senden, B., Nilsen, T., & Teig, N. (2023). The validity of student ratings of teaching quality: Factorial structure, comparability, and the relation to achievement. *Studies in Educational Evaluation*, *78*, 101274. https://doi.org/10.1016/j.stueduc.2023.101274
Sheldon, K. M., & Hilpert, J. C. (2012). The balanced measure of psychological needs (BMPN) scale: An alternative domain general measure of need satisfaction. *Motivation and Emotion, 36,* 439–451. https://doi.org/10.1007/s11031-012-9279-4
Silver, E. A., & Stein, M. K. (1996). The QUASAR project: The "revolution of the possible" in mathematics instructional reform in urban middle schools. *Urban Education, 30*(4), 476–521. https://doi.org/10.1177/0042085996030004006
Stallings, J. (1980). Allocated academic learning time revisited, or beyond time on task. *Educational Researcher, 9*(11), 11–16. https://doi.org/10.3102/0013189X009011011
Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics, 41*(5), 481–520. https://doi.org/10.3102/1076998616646200
Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488. https://doi.org/10.2307/1163292
Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50–80. https://doi.org/10.1080/ 1380361960020103
Styck, K. M., Anthony, C. J., Sandilos, L. E., & DiPerna, J. C. (2020). Examining rater effects on the classroom assessment scoring system. *Child Development, 92*(3), 976–993. https://doi.org/10.1111 /cdev.13460
Sun, H., & Chen, A. (2010). A pedagogical understanding of the self-determination theory in physical education. *Quest, 62*(4), 364–384. https://doi.org/10.1080/00336297.2010.10483655
Sun, Y., Liu, R.-D., Oei, T.-P., Zhen, R., Ding, Y., & Jiang, R. (2020). Perceived parental warmth and adolescents' math engagement in China: The mediating roles of need satisfaction and math self-efficacy. *Learning and Individual Differences, 78*, 101837. https://doi.org/10.1016/ j.lindif.2020.101837
Talić, I., Scherer, R., Marsh, H. W., Greiff, S., Möller, J., & Niepel, C. (2022). Uncovering everyday dynamics in students' perceptions of instructional quality with sampling. *Learning and Instruction*, *81*, 101594. https://doi.org/10.1016/j.learninstruc.2022.101594
Theis, D., Sauerwein, M., & Fischer, N. (2020). Perceived quality of instruction: The relationship among indicators of students' basic needs, mastery goals, and academic achievement. *British Journal of Educational Psychology*, *90*, 176–192. https://doi.org/10.1111/bjep.12313

Turner, J. C., & Meyer, D. K. (2004). A classroom perspective on the principle of moderate challenge in mathematics. *The Journal of Educational Research*, *97*(6), 311–318. https://psycnet.apa.org/doi/10.3200/JOER.97.6.311-318

Van der Kaap-Deeder, J., Soenens, B., Ryan, R. M., & Vansteenkiste, M. (2020). Manual of the Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS). Ghent University, Belgium.

van der Scheer, E. A., Bijlsma, H. J. E., & Glas, C. A. W. (2019). Validity and reliability ofstudent perceptions of teaching quality in primary education. *School Effectiveness and School Improvement, 30*(1), 30–50. https://doi.org/10.1080/09243453.2018.1539015

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. In T. C. Urdan & S. A. Karabenick (Eds.), *Advances in Motivation and Achievement: Vol. 16. The decade ahead: Theoretical perspectives on motivation and achievement* (1st ed., pp. 105–165). Emerald Group Publishing Limited. https://doi.org/10.1108/S0749-7423(2010)000016A007

Vieluf, S., Praetorius, A.-K., Rakoczy, K., Kleinknecht, M., & Pietsch, M. (2020). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts: Ein kritischer Vergleich verschiedener Modellvarianten [Opportunity-use model of effective teaching: A critical comparison of different models]. *Zeitschrift Für Pädagogik, 66*(1), 63–80. https://doi.org/10.3262/ZPB2001063

Vieluf, S. (2022). Wie, wann und warum nutzen Schüler*innen Lerngelegenheiten im Unterricht? Eine übergreifende Diskussion der Beiträge zum Thementeil. [How, when, and why do students use learning opportunities in the classroom? A comprehensive discussion of the contributions to the thematic section]. *Unterrichtswissenschaft*, *50*(2), 265-286. https://doi.org/10.1007/s42010-022-00144-z

Vieluf, S., & Klieme, E. (2023). Teaching effectiveness revisited through the lens of practice theories. In A-K. Praetorius & C. Y. Charalambous (Eds.), *Theorizing teaching: Current status and open issues* (pp. 57–95). Cham: Springer International Publishing.

Vu, T., Magis-Weinberg, L., Jansen, B. R., van Atteveldt, N., Janssen, T. W., Lee, N. C., ... & Meeter, M. (2022). Motivation-achievement cycles in learning: A literature review and research agenda. *Educational Psychology Review*, *34*(1), 39–71. https://doi.org/10.1007/s10648-021-09616-7

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, *108*(5), 705–721. https://doi.org/10.1037/edu0000075

Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction, 28*, 12–23. https://doi.org/10.1016/j.learninstruc.2013.04.002

Wang, Y., Tian, L., & Huebner, E. S. (2019). Basic psychological needs satisfaction at school, behavioral school engagement, and academic achievement: Longitudinal reciprocal relations among elementary school students. *Contemporary Educational Psychology*, *56*, 130–139. https://doi.org/10.1016/j.cedpsych.2019.01.003

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249–294. https://doi.org/10.3102/00346543063003249

Warner, G. J., Fay, D., & Spörer, N. (2017). Relations among personal initiative and the development of reading strategy knowledge and reading comprehension. *Frontline Learning Research*, *5*(2), 1–23. https://doi.org/10.14786/flr.v5i2.272

Wentzel, K. R., & Miele, D. B. (2016). *Handbook of motivation at school* (2nd ed.). Routledge.

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265–310. https://doi.org/10.1016/0273-2297(92)90011-P

Wigfield, A., & Eccles, J. (2002). *Development of achievement motivation*. Academic Press.

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique. *Organizational Research Methods*, *13*(3), 477–514.

https://doi.org/10.1177/1094428110366036

Zhang, T., Solmon, M. A., Kosma, M., Carson, R. L., & Gu, X. (2011). Need support, need satisfaction, intrinsic motivation, and physical activity participation among middle school students. *Journal of Teaching in Physical Education, 30*(1), 51–68. https://doi.org/10.1123/jtpe.30.1.51

Zee, M., Koomen, H. M. Y., & van der Veen, I. (2013). Student-teacher relationship quality and academic adjustment in upper elementary school: The role of student personality. *Journal of School Psychology*, *51*(4), 517–533. https://doi.org/10.1016/j.jsp.2013.05.003

Zhen, R., Liu, R.-D., Ding, Y., Wang, J., Liu, Y., & Le Xu (2017). The mediating roles of academic self-efficacy and academic emotions in the relation between basic psychological needs satisfaction and learning engagement among Chinese adolescent students. *Learning and Individual Differences, 54,* 210–216. https://doi.org/10.1016/j.lindif.2017.01.017

Zhu, Y., & Kaiser, G. (2022). Impacts of classroom teaching practices on students' mathematics learning interest, mathematics self-efficacy and mathematics test achievements: a secondary analysis of Shanghai data from the international video study Global Teaching Insights. *ZDM–Mathematics Education*, *54*(3), 581–593. https://doi.org/10.1007/s11858-022-01343-9

Zhou, J., Huebner, E. S., & Tian, L. (2021). The reciprocal relations among basic psychological need satisfaction at school, positivity and academic achievement in Chinese early adolescents. *Learning and Instruction, 71*, 101370. https://doi.org/10.1016/j.learninstruc.2020.101370

Ziegelbauer, S. (2009). *Denkprozesse lernwirksam anregen. Sensortechnik im modernen Physikunterricht.* [Stimulating thought processes in an effective way. Sensor technology in modern Physics lessons.] Tectum

SUPPLEMENTARY MATERIAL

Revisiting the Three Basic Dimensions model: A critical empirical investigation of the indirect effects of student-perceived teaching quality on student outcomes

1. Scales and items used in the study

   1.1. Items, reliabilities (ω), and descriptive statistics for the subscales

Items	ID-Pre	ω_{within}	$\omega_{between}$	*M*	*SD*
Cognitive activation with discourse (T1)	**sqa_cogdisc**	.65	.74	2.62	.49
Cognitive activation	**sqa_cogact**	.52	.85	2.60	.53
Our mathematics teacher presents tasks for which there is no obvious solution.	sqa18e	-	-	2.26	.88
Our mathematics teacher presents tasks that require us to apply what we have learned to new contexts.	sqa18f	-	-	3.09	.72
Our mathematics teacher gives tasks that require us to think critically.	sqa18g	-	-	2.53	.79
Our mathematics teacher asks us to decide on our own procedures for solving complex tasks.	sqa18h	-	-	2.53	.86
Discourse	**sqa_discourse**	.64	.87	2.65	.71
Our mathematics teacher gives us opportunities to explain our ideas.	sqa18i	-	-	3.13	.86
Our mathematics teacher encourages us to question and critique arguments made by other students.	sqa18j	-	-	2.73	.93
Our mathematics teacher requires us to engage in discussions among ourselves.	sqa18k	-	-	2.08	.96
Classroom management (T1)	**sqa_classman_mon_ removed**	.60	.95	3.01	.47
Disruptions	**sqa_disruptions**	.76	.99	2.79	.72
When the lesson begins, our mathematics teacher has to wait quite a long time for us to quieten down.	sqa20a_rec	-	-	2.83	.79
We lose quite a lot of time because of students interrupting the lesson.	sqa20b_rec	-	-	2.78	.83
There is much disruptive noise in this classroom.	sqa20c_rec	-	-	2.77	.85
Rule clarity	**sqa_rulecl**	-	-	2.92	.72
In our teacher's class, we are aware of what is allowed and what is not allowed.	sqa20d	-	-	3.37	.73
In our teacher's class, we know why certain rules are important	sqa20e	-	-	3.14	.74
Teachers managing disruptions	**sqa_tmd**	-	-	3.19	.70
Our teacher manages to stop disruptions quickly.	sqa20f	-	-	3.25	.75
Our teacher reacts to disruptions in such a way that the students stop disturbing learning.	sqa20g	-	-	3.14	.79
Transitions		-	-	-	-
In our teacher's class, transitions from one phase of the lesson to the other (e.g., from <class> discussions to individual work) take a lot of time.	sqa20h_rec	-	-	2.77	.76
Monitoring (removed from the analyses)	**sqa_monitoring**	-	-	2.78	.71

Our teacher is immediately aware of students doing something else.	sqa20i	-	-	2.77	.81
Our teacher is aware of what is happening in the classroom, even if he or she is busy with an individual student.	sqa20j	-	-	2.80	.80
Student support (T1)	**sqa_support_all**	.86	.98	2.98	.56
Teacher support for learning	**sqa_tesup**	.78	.97	3.03	.71
Our mathematics teacher gives extra help when we need it.	sqa21a	-	-	3.21	.77
Our mathematics teacher continues teaching until we understand.	sqa21b	-	-	2.96	.85
Our mathematics teacher helps us with our learning.	sqa21c	-	-	2.93	.82
Competence support	**sqa_supcom**	.82	.97	2.92	.72
Our mathematics teacher makes me feel confident in my ability to do well in the <course>.	sqa21d	-	-	2.85	.93
Our mathematics teacher listens to my view on how to do things.	sqa21e	-	-	2.99	.80
I feel that our mathematics teacher understands me.	sqa21f	-	-	2.93	.91
Our mathematics teacher makes me feel confident in my ability to learn the material.	sqa21g	-	-	2.93	.64
Autonomy support	**sqa_supaut**	.63	.82	2.98	.55
Our mathematics teacher provides me with different alternatives (e.g. learning materials or tasks).	sqa21h	-	-	2.67	.90
Our mathematics teacher encourages me to find the best way to proceed by myself.	sqa21i	-	-	2.83	.82
Our mathematics teacher lets me work on my own.	sqa21j	-	-	3.32	.67
Our mathematics teacher appreciates it when different solutions come up for discussion.	sqa21k			3.12	.76
Depth of processing (T1)	**sqa_usecogact**	.67	1.00	2.66	.63
I keep thinking about tasks until I really understand them.	sqa17d	-	-	2.96	.81
I think intensively about the mathematical content.	sqa17e	-	-	2.54	.80
I develop my own ideas regarding the topic taught.	sqa17f	-	-	2.47	.85
Time-on-task (T1)	**sqa_usetot**	.73	.95	3.10	.54
I pay attention in mathematics class.	sqa17j	-	-	3.19	.65
I listen to the instruction given in class.	sqa17k	-	-	3.33	.60
I let my mind wander during the lessons.	sqa17l_rec	-	-	2.76	.75
Need satisfaction (T1)	**sqa_useselfdet**	.63	.90	2.86	.64
I feel I can decide on things on my own. *(autonomy)*	sqa17g	-	-	2.50	.88
I feel understood by my mathematics teacher. *(relatedness)*	sqa17h	-	-	2.98	.90
I feel confident in my ability to learn this material. *(competence)*	sqa17i	-	-	3.11	.74
Interest (T1)	**sqa_interest**	.85	.98	2.40	.78
I am interested in mathematics.	sqa14a	-	-	2.73	.90
I often think that what we are talking about in my mathematics class is interesting.	sqa14b	-	-	2.43	.88
After mathematics class I am often already curious about the next mathematics class.	sqa14c	-	-	2.05	.86

Interest (T2)	**sqb_pint**	.81	.99	2.17	.73
I was interested in the topic of quadratic equations.	sqb03a	-	-	2.49	.88
I often thought that what we were talking about in my mathematics class during the unit on quadratic equations was interesting.	sqb03b	-	-	2.19	.84
After my mathematics class on the topic of quadratic equations I was often already curious about the next mathematics class.	sqb03c	-	-	1.83	.80

Note. For classroom management, subdimensions are determined according to the GTI technical report (Praetorius et al., pp. 2020b

1.2. The links to the TALIS Video Study

https://www.oecd.org/education/school/global-teaching-insights-technical-documents.htm

1.2.1. Student questionnaires

https://www.oecd.org/education/school/GTI-TechReport-AnnexD1.pdf

https://www.oecd.org/education/school/GTI-TechReport-AnnexD2.pdf

1.2.2. Student mathematics tests

https://www.oecd.org/education/school/GTI-TechReport-AnnexE2.pdf

https://www.oecd.org/education/school/GTI-TechReport-AnnexE3.pdf

2. R code for the analyses

2.1. R code for the preliminary analysis

2.1.1. Descriptive analysis

Descriptives for the variables

```
mean(mydata$sqa_cogdisc, na.rm=T)
sd(mydata$sqa_cogdisc, na.rm=T)

mean(mydata$sqa_classman_mon_removed, na.rm=T)
sd(mydata$sqa_classman_mon_removed, na.rm=T)

mean(mydata$sqa_support_all, na.rm=T)
sd(mydata$sqa_support_all, na.rm=T)

mean(mydata$sqa_usecogact, na.rm=T)
sd(mydata$sqa_usecogact, na.rm=T)

mean(mydata$sqa_usetot, na.rm=T)
sd(mydata$sqa_usetot, na.rm=T)
```

```
mean(mydata$sqa_useselfdet, na.rm=T)
sd(mydata$sqa_useselfdet, na.rm=T)

mean(mydata$sqa_interest, na.rm=T)
sd(mydata$sqa_interest, na.rm=T)

mean(mydata$sqa_ach, na.rm=T)
sd(mydata$sqa_ach, na.rm=T)

mean(mydata$sqb_pint, na.rm=T)
sd(mydata$sqb_pint, na.rm=T)

mean(mydata$sqb_ach, na.rm=T)
sd(mydata$sqb_ach, na.rm=T)
```

2.1.2 Bivariate correlations

```
mydata_3 <- data.frame(mydata$IDTEACHER, mydata$sqa_cogdisc,
        mydata$sqa_classman_mon_removed, mydata$sqa_support_all,
        mydata$sqa_usecogact, mydata$sqa_usetot, mydata$sqa_useselfdet,
        mydata$sqa_interest, mydata$sqa_ach, mydata$sqb_pint, mydata$sqb_ach)

correlation(mydata_3, multilevel = TRUE) #between level correlations
correlation(mydata_3) #within level correlations
```

2.1.3 Reliabilities

McDonald's Omega (ω)

```
# Cognitive activation
    omegaSEM(
     items = c("sqa18e", "sqa18f", "sqa18g", "sqa18h"),
     id = "IDTEACHER",
     data = mydata,
     savemodel = FALSE)

# Classroom management (monitoring removed)
    omegaSEM(
     items = c("sqa20a_rec", "sqa20b_rec", "sqa20c_rec", "sqa20d", "sqa20e", "sqa20f",
    "sqa20g", "sqa20h_rec"),
     id = "IDTEACHER",
     data = mydata,
     savemodel = FALSE)
```

```
# Student support
        omegaSEM(
         items = c("sqa21a", "sqa21b", "sqa21c", "sqa21d", "sqa21e", "sqa21f", "sqa21g",
        "sqa21h", "sqa21i", "sqa21j", "sqa21k"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)

# Depth of processing
        omegaSEM(
         items = c("sqa17d", "sqa17e", "sqa17f"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)

# Time-on-task
        omegaSEM(
         items = c("sqa17j", "sqa17k", "sqa17l_rec"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)

# Need satisfaction
        omegaSEM(
         items = c("sqa17g", "sqa17h", "sqa17i"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)

# Interest (T1)
        omegaSEM(
         items = c("sqa14a", "sqa14b", "sqa14c"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)

# Interest (T2)
        omegaSEM(
         items = c("sqb03a", "sqb03b", "sqb03c"),
         id = "IDTEACHER",
         data = mydata,
         savemodel = FALSE)
```

ICC1 and ICC2

```
multilevel.descript(mydata$sqa_cogdisc, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.103, ICC(2) = 0.717
multilevel.descript(mydata$sqa_classman_mon_removed, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.368, ICC(2) = 0.928
multilevel.descript(mydata$sqa_support_all, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.263, ICC(2) = 0.887
multilevel.descript(mydata$sqa_usecogact, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.043, ICC(2) = 0.496
multilevel.descript(mydata$sqa_usetot, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.036, ICC(2) = 0.450
multilevel.descript(mydata$sqa_useselfdet, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.111, ICC(2) = 0.734
multilevel.descript(mydata$sqa_interest, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.129, ICC(2) = 0.766
multilevel.descript(mydata$sqa_ach, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.204, ICC(2) = 0.850
multilevel.descript(mydata$sqb_pint, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.116, ICC(2) = 0.733
multilevel.descript(mydata$sqb_ach, cluster = mydata$IDTEACHER)
        # ICC(1) = 0.160, ICC(2) = 0.804
```

2.2. R code for the main analysis

2.2.1. Direct effects models

a. Cognitive activation – outcomes

```
Model_direct_cognitive_activation <- '
  level: 1

  sqb_ach ~ sqa_cogdisc
  sqb_pint ~ sqa_cogdisc

  sqb_ach ~ sqa_ach
  sqb_pint ~ sqa_interest

  sqa_cogdisc ~~ sqa_ach
  sqa_cogdisc ~~ sqa_interest
  sqa_ach ~~ sqa_interest
```

```
level: 2

sqb_ach ~ sqa_cogdisc
sqb_pint ~ sqa_cogdisc

sqb_ach ~ sqa_ach
sqb_pint ~ sqa_interest

sqa_cogdisc ~~ sqa_ach
sqa_cogdisc ~~ sqa_interest
sqa_ach ~~ sqa_interest
'
fitModel_direct_cognitive_activation <- sem(Model_direct_cognitive_activation, data = mydata,
cluster="IDTEACHER", estimator = "MLR")

summary(fitModel_direct_cognitive_activation, fit.measures = TRUE, standardized = TRUE, rsquare
= TRUE)
```

b. Classroom management – outcomes

```
Model_direct_classroom_management <- '
level: 1

sqb_ach ~ sqa_classman_mon_removed
sqb_pint ~ sqa_classman_mon_removed

sqb_ach ~ sqa_ach
sqb_pint ~ sqa_interest

sqa_classman_mon_removed  ~~ sqa_ach
sqa_classman_mon_removed  ~~ sqa_interest
sqa_ach ~~ sqa_interest

level: 2

sqb_ach ~ sqa_classman_mon_removed
sqb_pint ~ sqa_classman_mon_removed
```

```
    sqb_ach ~ sqa_ach
    sqb_pint ~ sqa_interest

    sqa_classman_mon_removed  ~~ sqa_ach
    sqa_classman_mon_removed  ~~ sqa_interest
    sqa_ach ~~ sqa_interest
    '
  fitModel_direct_classroom_management <- sem(Model_direct_classroom_management, data =
  mydata, cluster="IDTEACHER", estimator = "MLR")

  summary(fitModel_direct_classroom_management, fit.measures = TRUE, standardized = TRUE,
  rsquare = TRUE)
```

c. Student support – outcomes

```
Model_direct_student_support <- '
  level: 1

  sqb_ach ~ sqa_support_all
  sqb_pint ~ sqa_support_all

  sqb_ach ~ sqa_ach
  sqb_pint ~ sqa_interest

  sqa_support_all ~~ sqa_ach
  sqa_support_all ~~ sqa_interest
  sqa_ach ~~ sqa_interest

  level: 2

  sqb_ach ~ sqa_support_all
  sqb_pint ~ sqa_support_all

  sqb_ach ~ sqa_ach
  sqb_pint ~ sqa_interest
```

```
  sqa_support_all ~~ sqa_ach
  sqa_support_all ~~ sqa_interest
  sqa_ach ~~ sqa_interest
  '
fitModel_direct_student_support <- sem(Model_direct_student_support, data = mydata,
cluster="IDTEACHER", estimator = "MLR")

summary(fitModel_direct_student_support, fit.measures = TRUE, standardized = TRUE, rsquare =
TRUE)
```

2.2.2. The mediation models

a. Cognitive activation – mediators – outcomes

```
Model_indirect_cognitive_activation <- '
  level: 1
  sqa_usecogact  ~ a1*sqa_cogdisc
  sqa_usetot     ~ b1*sqa_cogdisc
  sqa_useselfdet ~ c1*sqa_cogdisc

  sqa_usecogact  ~ m1*sqa_interest + t1*sqa_ach
  sqa_usetot     ~ n1*sqa_interest + y1*sqa_ach
  sqa_useselfdet ~ o1*sqa_interest + z1*sqa_ach

  sqb_ach  ~  d1*sqa_usecogact + e1*sqa_usetot + f1*sqa_useselfdet
  sqb_pint  ~  g1*sqa_usecogact + h1*sqa_usetot + j1*sqa_useselfdet

  sqb_ach  ~ k1*sqa_cogdisc
  sqb_pint ~ l1*sqa_cogdisc

  sqb_ach ~ p1*sqa_ach
  sqb_pint ~ r1*sqa_interest

  sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
  sqa_useselfdet ~~ sqa_usetot

  sqa_ach ~~ sqa_cogdisc
  sqa_interest ~~ sqa_cogdisc
  sqa_ach ~~ sqa_interest

  a1d1 := a1*d1
  a1g1 := a1*g1

  b1e1 := b1*e1
```

```
b1h1 := b1*h1

c1f1  := c1*f1
c1j1  := c1*j1



level: 2
sqa_usecogact  ~ a2*sqa_cogdisc
sqa_usetot     ~ b2*sqa_cogdisc
sqa_useselfdet ~ c2*sqa_cogdisc

sqa_usecogact  ~ m2*sqa_interest + t2*sqa_ach
sqa_usetot     ~ n2*sqa_interest + y2*sqa_ach
sqa_useselfdet ~ o2*sqa_interest + z2*sqa_ach

sqb_ach  ~  d2*sqa_usecogact + e2*sqa_usetot + f2*sqa_useselfdet
sqb_pint  ~  g2*sqa_usecogact + h2*sqa_usetot + j2*sqa_useselfdet

sqb_ach  ~ k2*sqa_cogdisc
sqb_pint ~ l2*sqa_cogdisc

sqb_ach ~ p2*sqa_ach
sqb_pint ~ r2*sqa_interest

sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
sqa_useselfdet ~~ sqa_usetot

sqa_ach ~~ sqa_cogdisc
sqa_interest ~~ sqa_cogdisc
sqa_ach ~~ sqa_interest

a2d2 := a2*d2
a2g2 := a2*g2

b2e2 := b2*e2
b2h2 := b2*h2

c2f2  := c2*f2
c2j2  := c2*j2
'
fitModel_indirect_cognitive_activation <- sem(Model_indirect_cognitive_activation, data = mydata, cluster="IDTEACHER", bootstrap = 1000, estimator = "MLR")

summary(fitModel_indirect_cognitive_activation, fit.measures = TRUE, standardized = TRUE, rsquare = TRUE)
```

```
# Bootstrapping for mediating effects
boot.fit <- parameterEstimates(fitModel_indirect_cognitive_activation, boot.ci.type="bca.simple",
level=0.95, ci=TRUE, standardized = T)
boot.fit
```

b. *Classroom management – mediators – outcomes*

```
Model_indirect_classroom_management <- '

  level: 1
  sqa_usecogact  ~ a1*sqa_classman_mon_removed
  sqa_usetot     ~ b1*sqa_classman_mon_removed
  sqa_useselfdet ~ c1*sqa_classman_mon_removed

  sqa_usecogact  ~ m1*sqa_interest + t1*sqa_ach
  sqa_usetot     ~ n1*sqa_interest + y1*sqa_ach
  sqa_useselfdet ~ o1*sqa_interest + z1*sqa_ach

  sqb_ach  ~  d1*sqa_usecogact + e1*sqa_usetot + f1*sqa_useselfdet
  sqb_pint  ~  g1*sqa_usecogact + h1*sqa_usetot + j1*sqa_useselfdet

  sqb_ach  ~ sqa_classman_mon_removed
  sqb_pint ~ sqa_classman_mon_removed

  sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
  sqa_useselfdet ~~ sqa_usetot

  sqa_ach ~~ sqa_classman_mon_removed
  sqa_interest ~~ sqa_classman_mon_removed
  sqa_ach ~~ sqa_interest

  sqb_ach ~ sqa_ach
  sqb_pint ~ sqa_interest

  a1d1 := a1*d1
  a1g1 := a1*g1

  b1e1 := b1*e1
  b1h1 := b1*h1

  c1f1 := c1*f1
  c1j1 := c1*j1

level: 2
sqa_usecogact  ~ a2*sqa_classman_mon_removed
sqa_usetot     ~ b2*sqa_classman_mon_removed
sqa_useselfdet ~ c2*sqa_classman_mon_removed
```

```
sqa_usecogact  ~ m2*sqa_interest + t2*sqa_ach
sqa_usetot     ~ n2*sqa_interest + y2*sqa_ach
sqa_useselfdet ~ o2*sqa_interest + z2*sqa_ach

sqb_ach  ~  d2*sqa_usecogact + e2*sqa_usetot + f2*sqa_useselfdet
sqb_pint  ~  g2*sqa_usecogact + h2*sqa_usetot + j2*sqa_useselfdet

sqb_ach  ~ sqa_classman_mon_removed
sqb_pint ~ sqa_classman_mon_removed

sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
sqa_useselfdet ~~ sqa_usetot

sqa_ach ~~ sqa_classman_mon_removed
sqa_interest ~~ sqa_classman_mon_removed
sqa_ach ~~ sqa_interest

sqb_ach ~ sqa_ach
sqb_pint ~ sqa_interest

a2d2 := a2*d2
a2g2 := a2*g2

b2e2 := b2*e2
b2h2 := b2*h2

c2f2 := c2*f2
c2j2 := c2*j2
'
fitModel_indirect_classroom_management <- sem(Model_indirect_classroom_management, data =
mydata, cluster="IDTEACHER", bootstrap = 1000, estimator = "MLR")

summary(fitModel_indirect_classroom_management, fit.measures = TRUE, standardized = TRUE,
rsquare = TRUE)

# Bootstrapping for mediating effects
boot.fit <- parameterEstimates(fitModel_indirect_classroom_management, boot.ci.type="bca.simple",
level=0.95, ci=TRUE, standardized = T)
boot.fit
```

C. Student support – mediators – outcomes

```
Model_indirect_student_support <- '

level: 1
sqa_usecogact  ~ a1*sqa_support_all
sqa_usetot     ~ b1*sqa_support_all
sqa_useselfdet ~ c1*sqa_support_all

sqa_usecogact  ~ m1*sqa_interest + t1*sqa_ach
sqa_usetot     ~ n1*sqa_interest + y1*sqa_ach
sqa_useselfdet ~ o1*sqa_interest + z1*sqa_ach

sqb_ach  ~  d1*sqa_usecogact + e1*sqa_usetot + f1*sqa_useselfdet
sqb_pint  ~  g1*sqa_usecogact + h1*sqa_usetot + j1*sqa_useselfdet

sqb_ach  ~ sqa_support_all
sqb_pint ~ sqa_support_all

sqb_ach ~ sqa_ach
sqb_pint ~ sqa_interest

sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
sqa_useselfdet ~~ sqa_usetot

sqa_ach ~~ sqa_support_all
sqa_interest ~~ sqa_support_all
sqa_ach ~~ sqa_interest

a1d1 := a1*d1
a1g1 := a1*g1

b1e1 := b1*e1
b1h1 := b1*h1

c1f1 := c1*f1
c1j1 := c1*j1

level: 2
sqa_usecogact  ~ a2*sqa_support_all
sqa_usetot     ~ b2*sqa_support_all
sqa_useselfdet ~ c2*sqa_support_all

sqa_usecogact  ~ m2*sqa_interest + t2*sqa_ach
sqa_usetot     ~ n2*sqa_interest + y2*sqa_ach
sqa_useselfdet ~ o2*sqa_interest + z2*sqa_ach
```

```
sqb_ach  ~  d2*sqa_usecogact + e2*sqa_usetot + f2*sqa_useselfdet
sqb_pint  ~  g2*sqa_usecogact + h2*sqa_usetot + j2*sqa_useselfdet

sqb_ach  ~ sqa_support_all
sqb_pint ~ sqa_support_all

sqb_ach ~ sqa_ach
sqb_pint ~ sqa_interest

sqa_usecogact ~~ sqa_useselfdet + sqa_usetot
sqa_useselfdet ~~ sqa_usetot

sqa_ach ~~ sqa_support_all
sqa_interest ~~ sqa_support_all
sqa_ach ~~ sqa_interest

a2d2 := a2*d2
a2g2 := a2*g2

b2e2 := b2*e2
b2h2 := b2*h2

c2f2  := c2*f2
c2j2  := c2*j2
'
fitModel_indirect_student_support <- sem(Model_indirect_student_support, data = mydata,
cluster="IDTEACHER", bootstrap = 1000, estimator = "MLR")

summary(fitModel_indirect_student_support, fit.measures = TRUE, standardized = TRUE, rsquare =
TRUE)

# Bootstrapping for mediating effects
boot.fit <- parameterEstimates(fitModel_indirect_student_support, boot.ci.type="bca.simple",
level=0.95, ci=TRUE, standardized = T)
boot.fit
```

3. Additional Analyses

3.1. Multilevel path analyses – Direct effects models (correlational) (T1-T1)

 # Cognitive activation (T1) – Outcomes (T1)

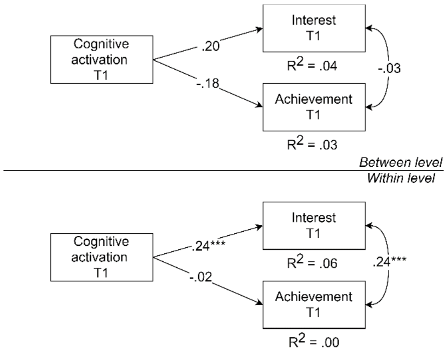

Figure S1. Multilevel correlational analysis for cognitive activation

 # Classroom management (T1) – Outcomes (T1)

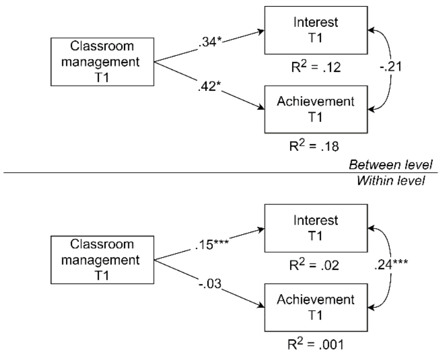

Figure S2. Multilevel correlational analysis for classroom management

 # Student support (T1) – Outcomes (T1)

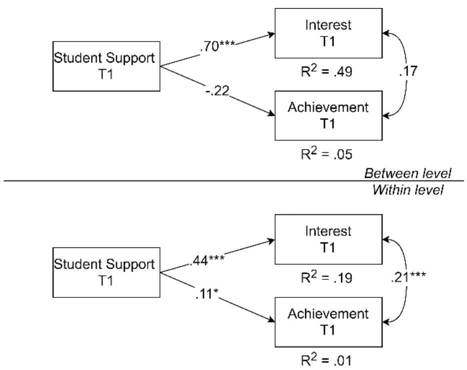

Figure S3. Multilevel correlational analysis for student support

3.2. Multilevel path analyses – Mediation models (correlational) (T1-T1-T1)

 # Cognitive activation (T1) – Mediators (T1) – Outcomes (T1)

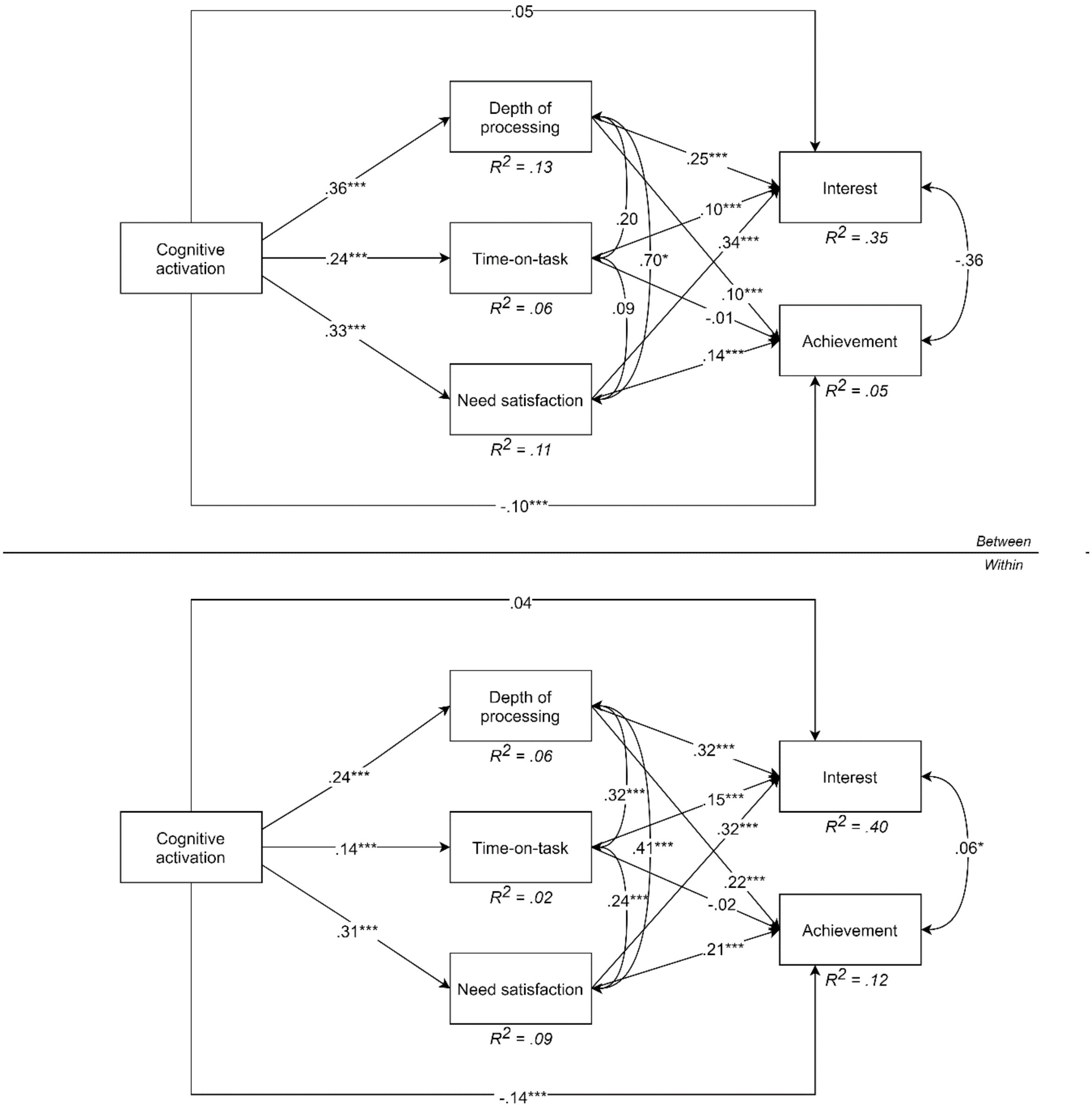

Figure S4. Multilevel correlational mediation analysis for cognitive activation

 # Classroom management (T1) – Mediators (T1) – Outcomes (T1)

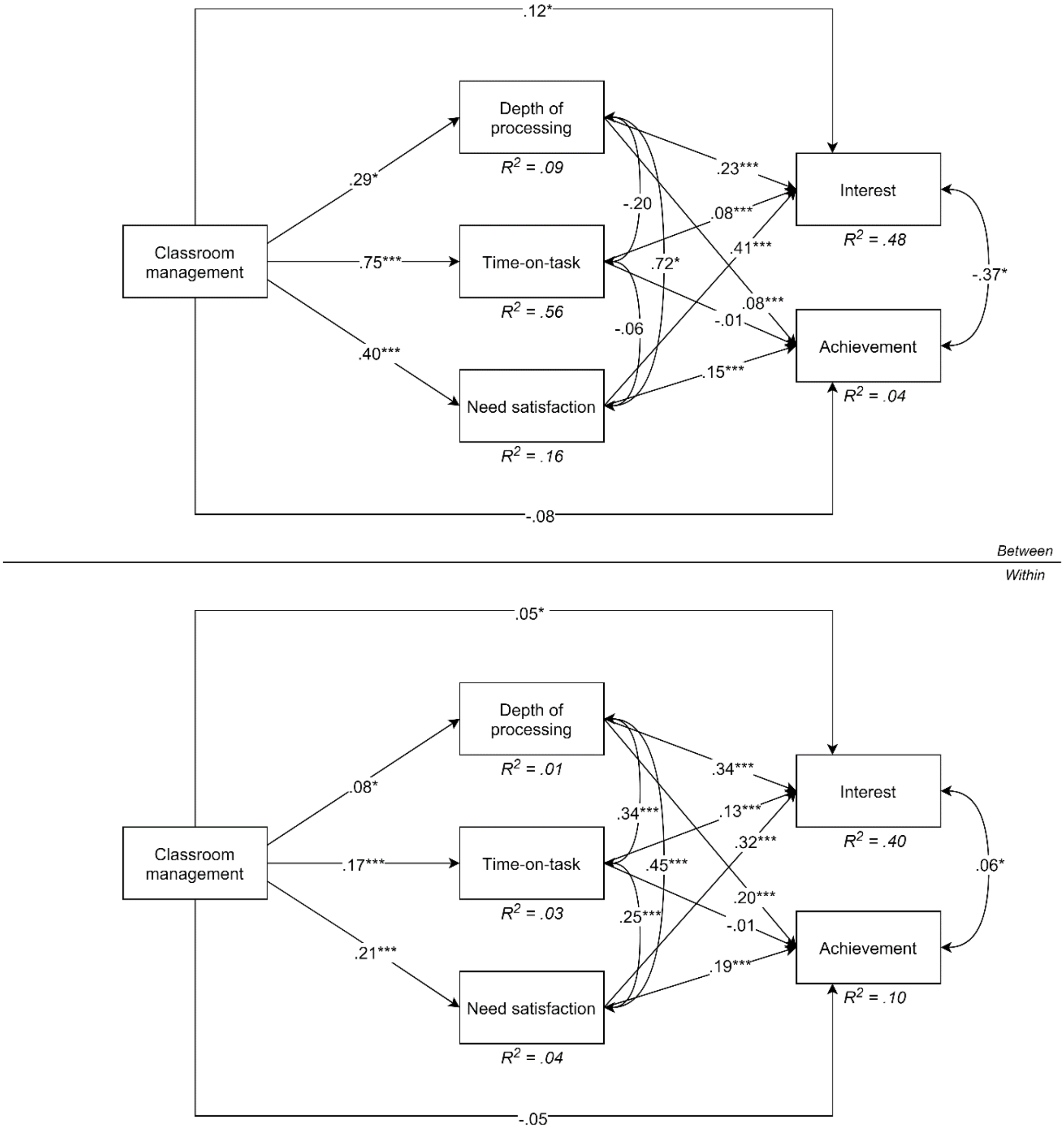

Figure S5. Multilevel correlational mediation analysis for classroom management

 # Student support (T1) – Mediators (T1) – Outcomes (T1)

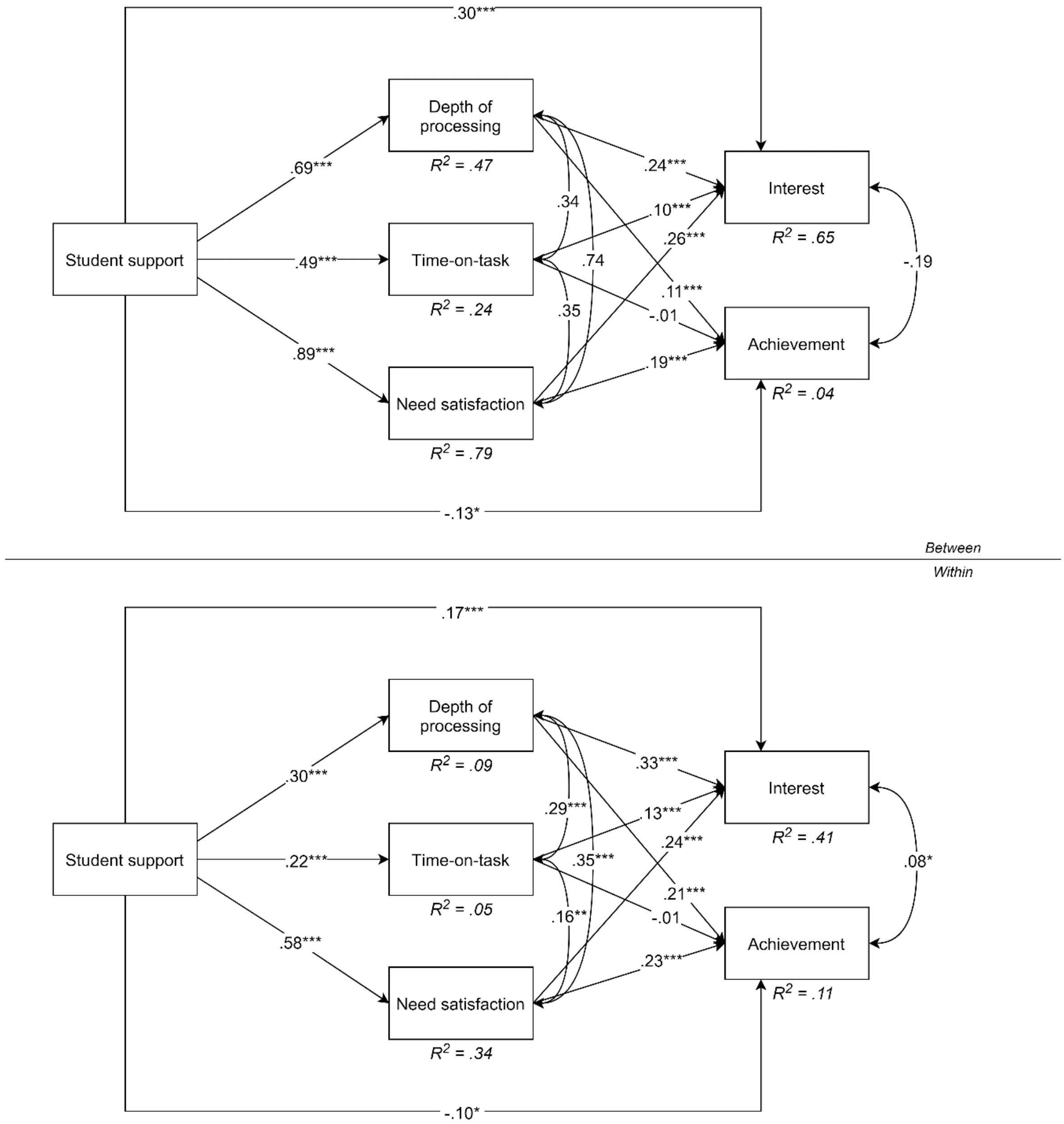

Figure S6. Multilevel correlational mediation analysis for student support

Table S1. Model fit indices for multilevel correlational path analyses (T1-T1-T1)

Model fit indices	χ^2	*p*	CFI	TLI	RMSEA [90%-CI]	$SRMR_{within}$	$SRMR_{between}$
Direct effects models							
1. Cognitive activation	15.569	–	.817	1.000	.000 [.000 – .000]	.000	.006
2. Classroom management	7.893	–	.920	1.000	.000 [.000 – .000]	.000	.002
3. Student support	7.454	–	.966	1.000	.000 [.000 – .000]	.000	.002
Mediation models							
4. Cognitive activation	984.582	< .000	.960	.891	.062 [.044 – .082]	.006	.172
5. Classroom management	39.342	< .000	.970	.918	.054 [.036 – .073]	.011	.164
6. Student support	36.621	< .000	.978	.939	.051 [.033 – .070]	.010	.199

Note. The SRMR at the between-level is related to the number of clusters in the dataset, in our case, there are 41 clusters. In cases with less than 200 clusters, the traditional 0.08 cutoff for SRMR may be too strict and does not necessarily indicate poor model fit (Asparouhov & Muthén, 2018, p. 13).

3.3. Number of classrooms vs. the number of days between T1 and T2

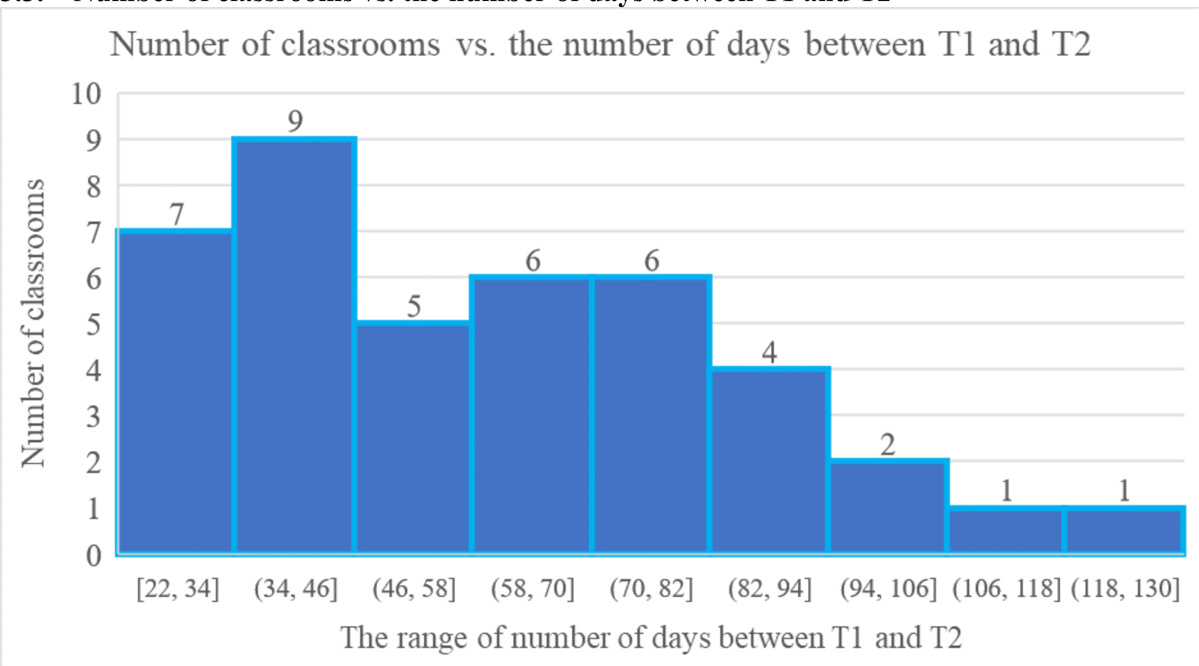

Figure S7. Number of classrooms vs. the number of days between T1 and T2