Chapter 3
The Relations Among School Climate, Instructional Quality, and Achievement Motivation in Mathematics

Ronny Scherer and Trude Nilsen

Abstract Instructional quality is considered to be an important classroom variable, as it is significantly related to student achievement and motivation in mathematics. Existing studies in educational effectiveness furthermore identified a positive relation between instructional quality and school climate, suggesting that the school environment plays a significant role in teachers' instructional practices. In order to bring together these two core findings, the relations among different aspects of school climate, instructional quality, and students' achievement motivation for the TIMSS 2011 grade eight mathematics data sets comprising 50 countries are investigated. In particular, the role of instructional quality as a potential mediator between school climate and student motivation is examined, thereby focusing on three aspects of school climate (emphasis on academic success, safety, and order in schools) and three aspects of achievement motivation (self-concept, intrinsic value, and extrinsic value). In general, there was a significant positive relation between instructional quality and achievement motivation at the classroom level in mathematics; in some countries, a partial mediation of instructional quality between school climate and achievement motivation was apparent. Four main patterns of relations occurred. These findings are discussed with respect to implications for educational effectiveness research.

Keywords School climate · Instructional quality · Student motivation · Two-level structural equation modeling mediation models · Trends in Mathematics and Science Study (TIMSS) 2011

R. Scherer (✉)
Faculty of Educational Sciences, Centre for Educational Measurement at the University of Oslo (CEMO), Oslo, Norway
e-mail: ronny.scherer@cemo.uio.no

T. Nilsen
Department of Teacher Education and School Research, University of Oslo, Oslo, Norway
e-mail: trude.nilsen@ils.uio.no

3.1 Rationale

Mathematics may be said to be at the heart of all science, technology, engineering, and mathematics (STEM) subjects. Motivating students to study these subjects is vital for a sustainable development within areas such as technology, economy, health, and environment. Yet, at the same time, an international concern for the decline of students' participation in STEM-related studies and careers has been raised (OECD 2014a). This concern seems to be rooted in a decrease of their STEM motivation (OECD 2014a). It is therefore important to motivate students for mathematics and for pursuing a career in science (Simpkins et al. 2006).

According to the widely accepted expectancy-value theory of achievement motivation, at least three motivational aspects are important for students' choices and performance in STEM areas: (a) self-beliefs, (b) intrinsic value, and (c) extrinsic value (Eccles and Wigfield 2002; Wigfield and Eccles 2000). Whereas (a) refers to students' beliefs in their capabilities, thereby reflecting their expectations of academic success, (b) and (c) are related to the subjective value assigned to subjects or tasks. Students may be driven by one or more of these aspects of motivation; either way, it is pertinent to identify factors that may promote these aspects of motivation and that lie within the power of the school.

Along with this challenge of identifying the motivational factors comes the question of how instruction and the school environment may contribute to student motivation in STEM subjects. According to previous research, teachers and their instruction matter more to student learning and motivation than any other school factor (Baumert et al. 2010; Creemers and Kyriakides 2008). The most important classroom variable is likely teachers' instructional quality, which affects both student achievement and motivation (Blömeke et al. 2013; Creemers and Kyriakides 2008; Fauth et al. 2014). Providing high quality instruction necessitates a safe and orderly school climate with a high priority for academic success (Thapa et al. 2013). Effective teaching is therefore challenged under conditions where teachers and students do not feel safe, where no order exists, and where academic success receives low priority. A healthy school climate consequently is important for student learning and motivation (Wang and Degol 2015).

In summary, a review of previous research indicates that, while instructional quality is important for student learning and motivation, school climate may contribute with ideal conditions for high quality instruction and hence promote learning and motivation. There are several aspects of school climate and motivation though, and the question of how instructional quality is related to these different aspects is complex (Good et al. 2009). This gap in research could be due to the extensive focus on achievement as a learning outcome as compared to motivation. Few studies have investigated the relations between school climate, instructional quality, and student motivation; these studies are almost exclusively focused on single-country analyses (Good et al. 2009; Seidel and Shavelson 2007; Wang and Degol 2015).

As a consequence, this chapter aims to address this research gap by investigating the relations among different aspects of school climate, instructional quality, and achievement motivation. Including all countries that participated in TIMSS 2011 provides a unique opportunity to investigate these relations across countries with widely different cultures from all continents.

3.2 Theoretical Framework

In this section, we first review the conceptualization of the three core constructs under investigation: school climate, instructional quality, and achievement motivation. We then present selected previous research on their relations.

3.2.1 School Climate

School climate is a broad concept that includes many dimensions (Thapa et al. 2013). Although it is defined somewhat differently across fields, certain key aspects have been found to be important to student learning. One of these aspects refers to academic climate, which is significantly positively related to student achievement and motivation (Wang and Degol 2015). Even though academic climate commonly refers to the extent to which learning and academic success is emphasized, there exists no consensus on its specific conceptualization. Hoy et al. (2006) referred to this aspect as academic optimism, a concept that reflects academic emphasis, collective efficacy, and faculty trust in parents and students. Together and individually, these three constructs have been found to be positively related to student learning (Goddard 2002; Hoy and Tschannen-Moran 1999; Hoy et al. 2006). Although the measurement of these constructs differed in further studies, this relation has been largely confirmed (Kythreotis et al. 2010; Martin et al. 2013; McGuigan and Hoy 2006; Nilsen and Gustafsson 2014),

In the context of TIMSS 2011, academic climate is represented and measured by the school emphasis on academic success (SEAS) scale. The underlying construct of SEAS has been found to be of great importance for students' learning outcomes and changes in performance across a number of countries (Martin et al. 2013; Nilsen and Gustafsson 2014). Conceptually, SEAS reflects the shared beliefs, capabilities, and trust among the members of the school institution (namely students, parents, teachers, and school leaders; Hoy et al. 2006; Nilsen and Gustafsson 2014). Among other aspects, SEAS comprises schools' trust in parents and students on the one hand and teachers' expectations for students' success on the other hand (Martin et al. 2013).

Another key aspect of school climate relates to a safe and orderly climate (Thapa et al. 2013; Wang and Degol 2015). Safety and order in schools refer to the degree of physical and emotional security, along with an orderly disciplinary climate

(Goldstein et al. 2008; Gregory et al. 2012; Wang and Degol 2015; Wilson 2004). Both safety and order are positively associated with student outcomes in a number of countries (Martin et al. 2013).

3.2.2 Instructional Quality

As detailed in Chap. 1, teachers' instructional quality comprise a number of aspects that have been shown to be highly important for student learning outcomes (Baumert et al. 2010; Creemers and Kyriakides 2008; Fauth et al. 2014; Good et al. 2009; Hattie 2009; Kunter et al. 2013; Seidel and Shavelson 2007). In the context of TIMSS 2011, students' ratings of instructional quality refer to aspects of supportive climate and clarity. We realize that this representation limits the rather broad concept of instructional quality to these two dimensions; yet, they are powerful indicators.

It is noteworthy that most studies investigate relations between instructional quality and achievement, while fewer include achievement motivation as an educational outcome (Fauth et al. 2014; Good et al. 2009).

3.2.3 Achievement Motivation

For decades, there has been an increasing concern for students' limited motivation in STEM subjects (NSF [US National Science Foundation] 2012; OECD 2007). Findings from TIMSS show that students' motivation for mathematics declines between grades four and eight. Moreover, previous studies have found significant differences with respect to the influence of gender on motivation in mathematics (Meece et al. 2006a, b), pointing to the necessity of accounting for gender in models of achievement motivation in mathematics. However, these findings vary across the different aspects of motivation (Wigfield et al. 2002).

According to the expectancy-value theory of achievement motivation proposed by Wigfield and Eccles (2000), there are some factors that directly influence student performance: expectation of success, interest-enjoyment value, attainment value, utility value, and cost. Wigfield and Eccles (2000) argued that students' expectations of success refer to their self-concept, reflected by the degree to which students believe they perform well in mathematics. The interest-enjoyment value refers to students' enjoyment and interest in a task or subject; Wigfield and Eccles (2000) claimed that this construct can be considered to be the intrinsic value or motivation of a subject or task (Deci and Ryan 1985; Harter 1981). Utility value refers to students' future career goals and aspirations, while attainment value reflects the personal importance of, for instance, mathematics. The last two factors reflect what is commonly referred to as 'extrinsic motivation'. In the current investigation, we will use the term extrinsic value. Costs reflect the negative aspects of motivation, such as performance anxiety and fear of both failure and success.

A number of studies have confirmed the importance of self-concept, intrinsic, and extrinsic value for students' career choices and performance (see Bandura 1997; Eccles and Wigfield 2002; Pintrich and Schunk 2002). As a consequence, we consider these aspects of achievement motivation as outcome variables in the current study.

3.3   Review of the Appropriate Level of Analysis

One question that arises with the measurement of the mentioned constructs concerns the appropriate level of analysis. In fact, given that both instructional quality and school climate are most often assessed using student or teacher ratings of the classroom or school environment, variation in these ratings may occur at different levels (namely student, classroom, school, or even country level; Klieme 2013). In order to make clear-cut decisions on the analysis level, a thorough review of the specific research questions is needed (Lüdtke et al. 2009). In the context of teacher effectiveness, research studying how the characteristics of the learning environment affect students' educational outcomes, such as their achievement, motivation, and self-beliefs, is the main focus (Klieme 2013; Lüdtke et al. 2009). Marsh et al. (2012) argued that the classroom or school level is the most appropriate in such scenarios. Nevertheless, as individual ratings of the learning environment may still vary and have a distinct meaning at the individual level, controlling for within-level variation is necessary (Lüdtke et al. 2009). One approach that has proven to be effective in modeling such situations refers to multilevel structural equation modeling (Scherer and Gustafsson 2015a).

In the current study, both instructional quality and school climate were assessed by individual ratings of students and teachers. We decided to study the relations among the two constructs and achievement motivation at the classroom level in a cross-country multi-group setting for two main reasons. First, the relation between instructional quality and student motivation clearly refers to a scenario in which the effects of the learning environment on student outcomes is the focus. Second, although teacher ratings of school climate may be considered a school-level construct, individual differences in these ratings still have a distinct meaning. In fact, teachers within a school may differ greatly in their perceptions of school climate, depending on their job satisfaction, well-being, status of professional development, and further individual-level factors (OECD 2014b; Wang and Degol 2015). Moreover, as teachers are the initiators of instructional practices in classrooms, their individual perceptions of the existing school climate are more important for their instruction than teachers' shared perceptions in a school. This finding has been supported by the results of the OECD Teaching and Learning International Survey (TALIS) 2013, which showed significant relations between individually perceived school climate and classroom instruction (OECD 2014b). As a consequence, we report the results at the classroom level.
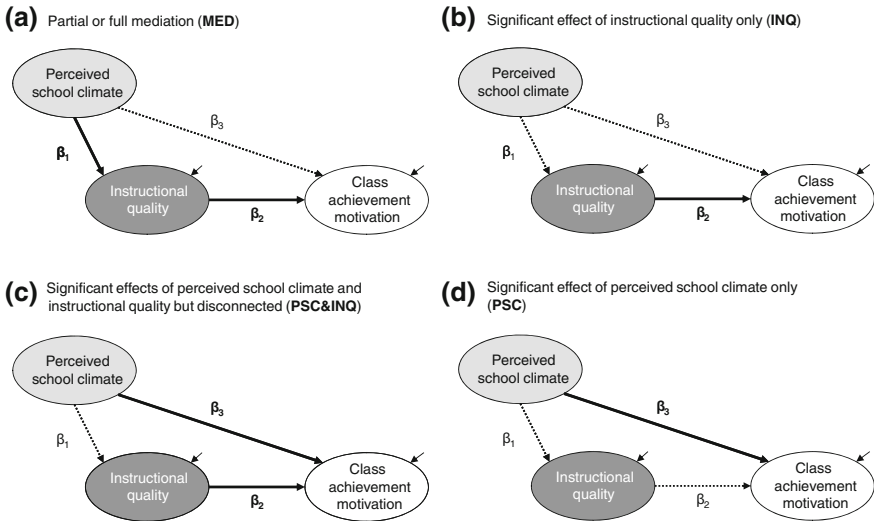
3.4   Research Model

Current frameworks for school effectiveness, such as the dynamic model of edu-
cational effectiveness, suggest that school climate influences both instructional
quality and learning outcomes (Creemers and Kyriakides 2010); it creates the
premise and foundation for instruction and learning (Thapa et al. 2013).

While only few studies have investigated the relations between school climate,
instructional quality, and achievement motivation, several studies have pointed
either to the importance of school climate for student motivation (Wang and Degol
2015) or to the importance of instructional quality for motivation (Fauth et al. 2014;
Wagner et al. 2015). However, rarely have all aspects of school climate been
investigated in concert (Wang and Degol 2015), and rarely have all aspects of
motivation been investigated in relation to school climate or instructional quality.
Hence, the relations between these concepts remain obscure.

On the basis of our literature review, we hypothesized that school climate,
instructional quality, and achievement motivation were related (Fig. 3.1). The
proposed relations at the classroom/teacher level are based on the core assumption
that a positively perceived school climate is a prerequisite for creating meaningful
instruction, which increases students' motivation to learn (Morin et al. 2014). To
account for this assumption, we examine the mediating role of instructional quality.
In this regard, we notice that instructional quality was measured by students' reports
(representing 'perceived instructional quality'), which were aggregated to the
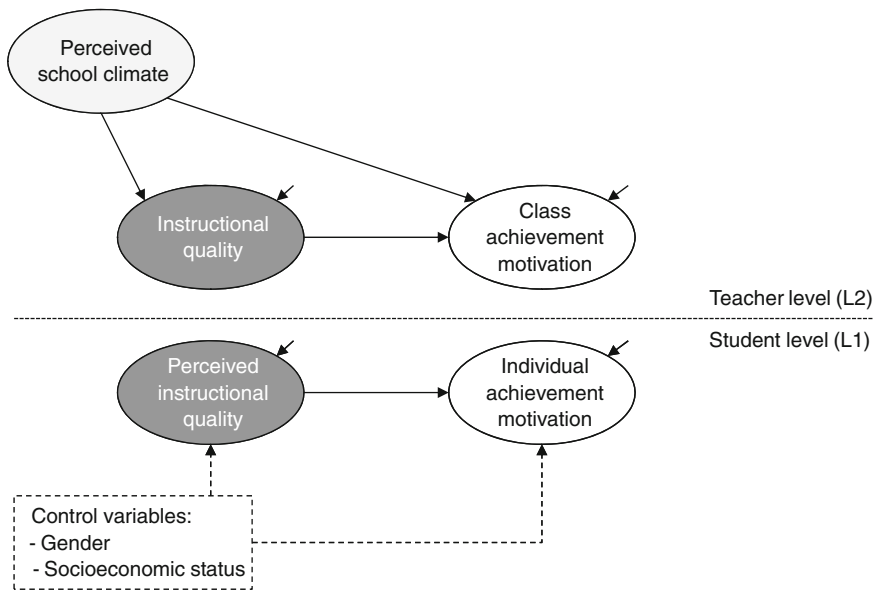
Fig. 3.1 Proposed research model, describing the relations among school climate, instructional
quality, and achievement motivation

Fig. 3.2 The four scenarios considered in the proposed research model. *Note* Dashed lines represent insignificant regression coefficients ($p > 0.05$). The relation between school climate and instructional quality is named β1, the relation between instructional quality and motivation is named β2, and the relation between school climate and motivation is named β3

classroom level ('instructional quality') by means of multilevel structural equation modeling.

However, due to gaps in previous research, the roles played by different aspects of school climate and achievement motivation are unclear. For instance, in some countries, SEAS may predict self-concept to a larger degree than other motivational constructs; at the same time, SEAS may be more important for extrinsic value in other countries. Different combinations of the aspects of school climate and motivation may therefore occur across countries. As a consequence, our research model considered three different aspects of school climate (i.e., SEAS, safety, and order in schools) and three aspects of achievement motivation (i.e., self-concept, intrinsic value, and extrinsic value).

As relationships between these factors and their statistical significance may also differ across countries, differing patterns of relations may result (see Fig. 3.2). These patterns of relations are described by four proposed model scenarios.

1. *Model MED*. The relation between teachers' perceived school climate and achievement motivation may be partially or fully mediated by instructional quality. In this scenario, there is a direct and significant link between school climate and instructional quality perceptions, suggesting that a positive and academically oriented school environment in which there is safety and order may foster a higher quality of instruction that motivates students and strengthens their self-beliefs and beliefs about the importance of mathematics for work and their future. This model proposes a (partial) mediation mechanism between the

three concepts of school climate, instructional quality, and achievement motivation (Deemer 2004).

2. *Model INQ*. There are contexts and situations in which only instructional quality is significantly related to achievement motivation; yet, perceived school climate has no effect. In these scenarios, perceived school climate and instruction are disconnected, indicating that the way teachers organize their instruction in mathematics classrooms is independent of how they perceive, for instance, an orientation toward academic success in the school environment.

3. *Model PSC&INQ*. In some countries, both perceived school climate and instructional quality are significantly related to motivational outcomes, although the two constructs are disconnected. On the one hand, this suggests that teacher perceptions of school climate do not influence the way in which they create learning environments in order to foster students' motivation and self-beliefs. On the other hand, besides the importance of instruction for the motivational outcomes, school climate directly relates to motivational outcomes.

4. *Model PSC*. In some countries and cultures, only perceived school climate, which is an indicator of the actual climate in schools, is significantly related to achievement motivation. In this scenario, school climate seems to be the dominant factor influencing student motivation. In fact, in some countries, a climate of safety and order in schools is considered to be a crucial prerequisite for student learning (Klieme et al. 2009; Mitchell and Bradshaw 2013). Moreover, existing research has suggested that school emphasis on academic success is significantly related to student performance in not only mathematics, but also science achievement tests across almost all TIMSS 2011 participating countries (Mullis et al. 2012; Nilsen and Gustafsson 2014). The relation between perceived school climate and students' motivational outcomes may not necessarily be positive; in fact, negative relations are also likely, particularly with SEAS. More specifically, a strong emphasis on academic success can create highly competitive learning environments that decrease students' motivation and self-beliefs due to a strong performance orientation (Chen and Vazsonyi 2013; Meece et al. 2006a, b). However, this negative relation is not limited to this model.

Our list of scenarios is not exhaustive; further potential scenarios, such as a model in which all relations are insignificant, or a model in which there is only a significant relation between school climate and instructional quality without any connection to achievement motivation, may occur. However, as these last scenarios occurred in only two cases and were of limited substantive relevance, their interpretation was limited.

In light of our considerations, we posed the following research question based on our proposed research model:

To what extent do the different scenarios in the proposed research model, as representatives of different patterns in the relations among school climate, instructional quality, and achievement motivation in mathematics, exist at the teacher/classroom level across the 50 participating TIMSS 2011 grade eight countries (Fig. 3.2)?

3.5 Method

3.5.1 Sample

The total TIMSS 2011 grade eight mathematics student sample, together with their teachers, formed the basis for the present study.[1] This sample comprised $n = 284{,}899$ students in 12,950 classrooms with an average classroom size of 22 students. All 50 participating TIMSS 2011 countries were included in the analyses. For further details on the sample, please refer to the TIMSS 2011 International Report (Mullis et al. 2012).

3.5.2 Measures

To assess the relations among school climate, instructional quality, and student achievement motivation, we used the existing TIMSS 2011 grade eight student and teacher scales to represent each of the three constructs. These scales and their conceptual underpinnings are briefly described here. For a detailed description of these measures, we refer the reader to the TIMSS 2011 Assessment Frameworks Report (Mullis et al. 2009).

School Climate

In order to capture different aspects of school climate, we followed Wang and Degol's (2015) systematic review and chose three scales as proxies for the construct: School emphasis on academic success, Safety, and Order in schools. Teachers had to rate a number of statements on a four-point agreement scale (0 = I disagree a lot, to 3 = I agree a lot).

SEAS: Teachers' ratings formed the basis for creating a latent variable for SEAS (Martin et al. 2013), because teachers are closer to the classrooms and students than principals. In the teacher questionnaire, teachers were asked to characterize the following within their school: teachers' understanding of and success in implementing the school's curriculum, teachers' expectations for student achievement, parental support for student achievement, and students' desire to do well in school. Hence, the teachers rated SEAS as an aspect of school climate in their schools.

---

[1]In Botswana, Honduras, and South Africa, the TIMSS 2011 study was conducted in grade nine.

Safety in Schools: Teachers' perceptions of safety in schools were indicated by three items ("This school is located in a safe neighborhood", "I feel safe at this school", and "This school's security policies and practices are sufficient").

Order in Schools: Teachers had to evaluate the degree of order and respect in their schools ("The students behave in an orderly manner", "The students are respectful of the teachers").

Instructional Quality

To measure aspects of instructional quality, we chose four items from the 'students engaged in mathematics lessons' scale. One item was removed (namely, "I think of things not related to the lesson"), as it was negatively worded and therefore referred to students' boredom and inattention rather than their perceptions of whether or not teachers engage them; the measurement model based on five items indicated a poor model fit across countries. The remaining four items referred to the clarity of teaching ("I know what my teacher expects me to do", "My teacher is easy to understand"), and the degree to which the teacher engages students to learn mathematics ("I am interested in what my teacher says", "My teacher gives me interesting things to do").[2] Students had to indicate their agreement with these statements on a four-point scale (0 = I disagree a lot, to 3 = I agree a lot). These items provided valid representations of instruction aimed at engaging students in learning (Scherer and Gustafsson 2015a). These four items were used as indicators of instructional quality in all analyses presented in this chapter. Moreover, as this chapter is not concerned with teachers' perceptions of their instruction, instructional quality was measured at the student level and subsequently aggregated to the classroom level.

Motivation

Measures of student motivation were retrieved from the student questionnaire scales. These scales were developed on the basis of expectancy-value theory of achievement motivation and referred to two main components: (a) students' expectations of success in mathematics, and (b) students' subjective task values (Wigfield and Eccles 2000). The first were indicated by ability beliefs (self-concept); the second were indicated by students' intrinsic and extrinsic values. Students had to indicate the degree to which they agreed to a number of statements

---

[2]This aspect of instructional quality is closely related to teachers' motivational support aimed at engaging students to learn in their mathematics lessons.

(0 = I disagree a lot, 1 = I disagree a little, 2 = I agree a little, 3 = I agree a lot). Since methodological research has clearly indicated that a mixture between positively and negatively worded items creates construct-irrelevant multidimensionality in assessments of motivational constructs such as self-concept (see Morin et al. 2015), we decided to omit negatively worded items, as these, by and large, measure a substantively different construct than positively worded items (Marsh and Gouvernet 1989). This has been confirmed for a number of measures that relied on self-ratings (see for example, Davison and Srichantra 1988; Greenberger et al. 2003; Marsh and Gouvernet 1989; Podsakoff et al. 2003; Preckel 2014; Scherer and Gustafsson 2015b). Van Sonderen et al. (2013) further pointed out that simply recoding reversely coded items does not solve the issue of the method bias created by negatively worded items; the hope to correct for potential response bias by introducing such items has not been fulfilled. There is evidence of this method bias in the measurement of motivational constructs in TIMSS (see Bofah and Hannula 2015; Marsh et al. 2013).

Self-concept: Students' self-concept in mathematics was originally assessed by seven items corresponding to the TIMSS 2011 'students confident in mathematics' scale, four of which were negatively worded; as decided, these latter items were deleted. The resultant scale comprised three items: "I usually do well in mathematics", "I learn things quickly in mathematics", and "I am good at working out difficult problems in mathematics". Although this decision limited the overall number of indicators of self-concept, existing research has shown that a three-item scale still provides a reliable and valid measure of students' self-concept (Gogol et al. 2014).

Intrinsic Value: In order to represent the intrinsic task value, we used the TIMSS 2011 'students like learning mathematics' scale, which comprised five items.[3] This scale refers to students' enjoyment and interest in learning mathematics (for example, "I enjoy learning mathematics").

Extrinsic Value: This value component of achievement motivation was represented by the 'students value mathematics' scale. The scale comprises differing aspects of the utility and attainment value of learning mathematics and its personal importance: "I think learning mathematics will help me in my daily life", "I need mathematics to learn other school subjects", "I need to do well in mathematics to get into the < university > of my choice", "I need to do well in mathematics to get the job I want", and "I would like a job that involves using mathematics".

---

[3]The original scale comprised six items, one of which was negatively formulated. As argued for the measurement of self-concept, we deleted this item to avoid method bias and construct-irrelevant multidimensionality.

Control Variables: Socioeconomic Status

To represent students' socioeconomic status, a variable derived from several items in the student questionnaire (students' ratings of the number of books at home, their parents' highest education and home study supports such as students having their own room and internet connection) was available in the TIMSS 2011 data set (the Home Educational Resources scale). We used the person estimate derived from a partial credit model in the TIMSS 2011 scaling procedure (Martin and Mullis 2012).

Control Variables: Gender

Gender served as another student-level covariate, because some research has suggested that gender differences may exist in student ratings for both achievement motivation and instructional quality (Lazarides and Ittel 2012; Meece et al. 2006a, b; Wigfield et al. 2002).

3.5.3 Statistical Analysis

We conducted a number of modeling steps comprising measurement invariance testing and multilevel structural equation modeling. In all analyses, robust maximum likelihood estimation (MLR) was used, with standard errors and tests of fit that were robust against non-normality of observations and the use of categorical variables in the presence of at least four response categories (Beauducel and Herzberg 2006; Rhemtulla et al. 2012).

Step 1 Measurement invariance testing

We applied multi-group confirmatory factor analysis (MGCFA) to test the measurement models of each construct included in the proposed research model for invariance across the 50 participating TIMSS 2011 countries. This step was necessary to ensure that the measures were, to a sufficient degree, comparable and to exclude measurement bias as a potential source of cross-country differences (Rutkowski and Svetina 2014). As instructional quality and the motivational constructs were measured at both the student and the classroom level, we tested for measurement invariance at these two levels by conducting (a) single-level MGCFA, and (b) multilevel MGCFA. For the latter, the student (individual) level was saturated, assuming only correlations among all items of a scale (Ryu 2014). For the school climate constructs, however, only (b) applied, because they were measured by teacher ratings.

Testing for measurement invariance, we specified a configural model, in which the number of factors and the pattern specified in the loading matrices were equal

across countries. Building upon this model, metric invariance additionally constrained the factor loadings to equality. Finally, scalar invariance assumed that the item intercepts were equal across countries in addition to the factor loadings. To ensure that the relations among the latent variables proposed in our research model were comparable across countries, at least metric invariance must hold (Millsap 2011). We evaluated these three invariance models with respect to their overall goodness-of-fit, and the changes in the goodness-of-fit statistics after introducing constraints on factor loadings and item intercepts. The configural model formed the basis for evaluating these changes.

To evaluate the changes in model fit, we followed the recommendations given by Rutkowski and Svetina (2014) and considered the changes of the incremental fit indices as practically insignificant if changes in the comparative fit index (CFI) were less than 0.020, and the root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR) changed by less than 0.020,[4] compared to the configural invariance model. These statistics are particularly sensitive to deviations from the invariance of factor loadings and intercepts (Chen 2007). Although these guidelines have been studied and applied in various contexts, we consider them to be only rough guidelines in the current investigation, as the number of groups is extraordinarily high.

Step 2 Multilevel structural equation modeling (MSEM)

On the basis of the measurement invariance testing results, we applied multilevel structural equation modeling to the data of each of the 50 countries and used the factor loadings obtained from the first step of invariance testing at the classroom level as fixed parameters in the measurement models of the constructs under investigation. This procedure may circumvent convergence problems and provides results in reasonable estimation time. Although we are aware that this fixed parameters country-by-country approach may result in less precise parameter estimates than a multi-group multilevel modeling approach that estimates the factor loadings across countries, there were significant advantages in reducing the number of model parameters and therefore simplifying the model estimation. In fact, according to our research model, in a multi-group MSEM describing the relations among, for instance, an orderly school climate, instructional quality, and academic self-concept in mathematics, more than 1500 parameters had to be estimated; the fixed parameters approach resulted in 36 estimated parameters per country. These figures illustrate the substantial reduction in estimation effort and model complexity gained by this approach, with only limited loss in precision.

In all country-by-country MSEM analyses, we tested the indirect effect of school climate on achievement motivation via instructional quality against zero to check

---

[4]Please note that for large numbers of groups (20 or more) more liberal criteria for the ΔRMSEA and ΔSRMR may be applied. In this sense, an increase in the RMSEA of less or equal than 0.030 could still be considered acceptable.

for potential (partial) mediation. The corresponding standard errors were obtained from M*plus* using an asymptotic estimation procedure (Muthén and Muthén 1998–2014; Preacher et al. 2010). We used an implicit latent group-mean centering and level-specific standardization (Stancel-Piątak and Desa 2014).[5]

In all analyses, missing data were handled using the full-information maximum likelihood procedure under the assumption that missing data occurred at random (Enders 2010). We furthermore included the mathematics teachers' weights in the analyses (MATWGT) to account for the sampling design applied in TIMSS 2011. The IDB (International Database) analyzer (IEA 2012) was used to prepare and merge the data. Significance testing was performed at the 5 % level.

3.6   Results

3.6.1   *Measurement Invariance Testing*

As already mentioned, we tested for measurement invariance for the constructs that play a role in our proposed research model (Fig. 3.1) in order to obtain evidence on sufficient degrees of comparability of measures across the 50 participating TIMSS 2011 countries. In this respect, we tested for invariance of the measurement models at the between (classroom/teacher) level for all constructs, and for invariance at the within (student) level for constructs based on student ratings.

The resulting goodness-of-fit statistics and the corresponding model comparisons indicated that, for the different aspects of achievement motivation and student ratings of instructional quality, both student- and classroom-level metric invariance could be established (see Appendix C). For the school climate scales capturing safety and order in schools, metric invariance can be assumed. Changes in goodness-of-fit indices exceeded the suggested cut-offs in only few instances, however: (a) the suggested cut-offs have not yet been evaluated in multi-group multilevel situations with a larger number of groups (in our case, countries) and can therefore only be regarded as approximate marking points; (b) the metric invariance model fitted the data reasonably well; and (c) while the CFI was substantially lower in the metric model, the TLI improved compared to the configural model, suggesting that there was mixed evidence on changes in these fit statistics. As a consequence, we accepted the metric invariance models for all constructs and levels and interpreted the invariance testing results as evidence for a sufficient degree of comparability. We therefore proceeded with comparing the relations among school climate, instructional quality, and achievement motivation across countries.

---

[5]The resulting standardized regression coefficients are those reported for the classroom/teacher level (in contrast to reporting the contextual or compositional effects; Marsh et al. 2012).

3.6.2 Multilevel Structural Equation Modeling

Since specifying our research model for three aspects of school climate (namely, SEAS, safety, and order in schools) and three aspects of achievement motivation (self-concept, intrinsic, and extrinsic value) has resulted in nine models and therefore a rich amount of data, we systematized the findings using the classification presented previously (see Fig. 3.2). Specifically, for each, we allocated the data of a particular country to one of our models: MED, INQ, PSC, and PSC&INQ. We here present detailed results for one of these models (results of the other models are provided in Appendix C). The goodness-of-fit statistics for the country-by-country MSEM analysis with fixed factor loadings in the measurement models of the constructs (see Sect. 3.5) were largely acceptable; in some cases, the statistics approached the cut-off value (CFI and TLI close to 0.90, RMSEA close to 0.08).

We studied the model for SEAS as an indicator of teachers' perceived school climate and students' intrinsic value. This model provided regression coefficients for the 50 countries (Table 3.1). On the basis of the direct and indirect effects, each country was assigned to one of the potential models.

Twelve countries fitted the MED model, where it was apparent that the SEAS-intrinsic value relation was at least partially mediated by instructional quality (indirect effect $\beta_1 \times \beta_2$: $M = 0.206$, $SD = 0.048$, $Mdn = 0.186$, $Min = 0.151$, $Max = 0.283$); for the South African data set, the mediation was negative due to a negative relation between SEAS and instructional quality. Twenty-six countries satisfied model INQ; the average path coefficient of instructional quality on students' intrinsic value was 0.844 ($SD = 0.087$, $Mdn = 0.861$, $Min = 0.590$, $Max = 0.985$). The remaining 11 countries fulfilled the model PSC&INQ, where both SEAS ($M[\beta_3] = 0.177$, $SD = 0.068$, $Mdn = 0.162$, $Min = 0.101$, $Max = 0.337$) and instructional quality ($M[\beta_2] = 0.865$, $SD = 0.040$, $Mdn = 0.865$, $Min = 0.787$, $Max = 0.933$) had significant effects on intrinsic value; in Turkey there was a negative relationship between SEAS and instructional quality. None of the countries fitted model PSC.

We were thus able to identify three out of the four proposed scenarios in our research. Interestingly, for the majority of countries, instructional quality was strongly associated with students' intrinsic value; for some countries, the relation between SEAS and intrinsic value was at least partially mediated via instructional quality. The latter result points to the existence of a potential mechanism among the three constructs.

We studied the proposed research model for each of the school climate aspects and aspects of achievement motivation and assigned them the appropriate model (Table 3.2).

Given the rich results, we here only highlight selected patterns. First, the results for each country are relatively consistent; the majority of countries display similar patterns across all aspects of achievement motivation, given a specific aspect of school climate. For example, the Australian data set indicated that the school climate-instructional quality-achievement motivation relation was mediated for

Table 3.1 Standardized direct and indirect effects in the model with SEAS as the school climate aspect and intrinsic value as the motivational outcome variable (see also Fig. 3.2 for explanation of scenarios)

Country	Direct effects			Indirect effect	Model
	β_1 (SE)	β_2 (SE)	β_3 (SE)	$\beta_1 \times \beta_2$ (SE)	
Armenia	–0.131 (0.098)	0.750 (0.059)[*]	0.118 (0.090)	−0.098 (0.073)	INQ
Australia	0.221 (0.066)[*]	0.827 (0.034)[*]	0.014 (0.054)	0.183 (0.055)[*]	MED
Bahrain	–0.226 (0.128)	0.855 (0.093)[*]	0.337 (0.092)[*]	–0.193 (0.120)	PSC&INQ
Chile	–0.003 (0.100)	0.897 (0.0039)[*]	0.070 (0.080)	–0.003 (0.090)	INQ
Chinese Taipei	–0.039 (0.091)	0.933 (0.037)[*]	0.188 (0.077)[*]	–0.037 (0.085)	PSC&INQ
England	0.173 (0.091)	0.836 (0.037)[*]	0.044 (0.058)	0.144 (0.076)	INQ
Finland	0.186 (0.090)[*]	0.902 (0.034)[*]	0.073 (0.063)	0.168 (0.081)[*]	MED
Georgia	0.096 (0.097)	0.859 (0.043)[*]	0.160 (0.069)[*]	0.083 (0.083)	PSC&INQ
Ghana	0.048 (0.112)	0.817 (0.090)[*]	0.061 (0.091)	0.039 (0.092)	INQ
Hong Kong SAR	0.252 (0.125)[*]	0.746 (0.064)[*]	0.277 (0.078)[*]	0.188 (0.089)[*]	MED
Hungary	0.076 (0.090)	0.916 (0.023)[*]	0.011 (0.053)	0.070 (0.082)	INQ
Indonesia	–0.051 (0.110)	0.943 (0.034)[*]	0.027 (0.061)	–0.049 (0.104)	INQ
Iran, Islamic Rep. of	–0.101 (0.087)	0.845 (0.040)[*]	0.238 (0.079)[*]	–0.086 (0.075)	PSC&INQ
Israel	–0.062 (0.078)	0.827 (0.042)[*]	0.147 (0.067)[*]	–0.051 (0.065)	PSC&INQ
Italy	0.048 (0.102)	0.985 (0.028)[*]	0.059 (0.057)	0.047 (0.100)	INQ
Japan	0.288 (0.090)[*]	0.895 (0.044)[*]	0.064 (0.073)	0.257 (0.078)[*]	MED
Jordan	0.071 (0.089)	0.963 (0.042)[*]	0.016 (0.064)	0.068 (0.086)	INQ
Kazakhstan	0.049 (0.103)	0.892 (0.034)[*]	0.130 (0.058)[*]	0.043 (0.092)	PSC&INQ
Korea, Rep. of	0.137 (0.072)	0.715 (0.048)[*]	–0.027 (0.060)	0.098 (0.053)	INQ
Lebanon	–0.095 (0.109)	0.787 (0.051)[*]	0.180 (0.083)[*]	–0.075 (0.087)	PSC&INQ

(continued)

Table 3.1 (continued)

Country	Direct effects			Indirect effect	Model
	β_1 (SE)	β_2 (SE)	β_3 (SE)	$\beta_1 \times \beta_2$ (SE)	
Lithuania	0.069 (0.082)	0.933 (0.022)*	0.044 (0.050)	0.065 (0.076)	INQ
Macedonia	0.081 (0.100)	0.878 (0.030)*	−0.035 (0.057)	0.071 (0.088)	INQ
Malaysia	0.105 (0.081)	0.883 (0.032)*	0.127 (0.048)*	0.092 (0.070)	PSC&INQ
Morocco	0.073 (0.095)	0.906 (0.063)*	0.091 (0.061)	0.066 (0.084)	INQ
New Zealand	0.206 (0.064)*	0.858 (0.030)*	−0.064 (0.062)	0.176 (0.056)*	MED
Norway	0.323 (0.105)*	0.866 (0.051)*	0.188 (0.067)*	0.280 (0.084)*	MED
Oman	0.215 (0.085)*	0.775 (0.034)*	0.103 (0.057)	0.166 (0.064)*	MED
Palestinian Nat'l Auth.	0.211 (0.084)*	0.780 (0.047)*	0.015 (0.067)	0.165 (0.067)*	MED
Qatar	0.190 (0.099)	0.860 (0.055)*	0.002 (0.082)	0.164 (0.089)	INQ
Romania	0.061 (0.079)	0.886 (0.039)*	−0.021 (0.054)	0.054 (0.070)	INQ
Russian Federation	0.214 (0.095)*	0.917 (0.028)*	0.007 (0.053)	0.197 (0.088)*	MED
Saudi Arabia	0.052 (0.114)	0.892 (0.033)*	−0.024 (0.062)	0.046 (0.102)	INQ
Singapore	0.066 (0.070)	0.803 (0.059)*	0.056 (0.073)	0.053 (0.057)	INQ
Slovenia	0.035 (0.069)	0.893 (0.031)*	0.163 (0.058)*	0.031 (0.061)	PSC&INQ
Sweden	0.206 (0.074)*	0.734 (0.050)*	0.135 (0.076)	0.151 (0.054)*	MED
Syrian Arab Rep.	−0.008 (0.124)	0.771 (0.065)*	0.006 (0.095)	−0.006 (0.096)	INQ
Thailand	−0.057 (0.112)	0.896 (0.045)*	0.087 (0.065)	−0.0051 (0.101)	INQ
Tunisia	0.002 (0.096)	0.881 (0.036)*	0.072 (0.071)	0.002 (0.084)	INQ
Turkey	−0.063 (0.091)	0.869 (0.031)*	−0.106 (0.053)*	−0.055 (0.078)	PSC(−)&INQ
Ukraine	0.266 (0.118)*	0.952 (0.027)*	0.037 (0.059)	0.253 (0.111)*	MED
United Arab Emirates	−0.098 (0.055)	0.834 (0.024)*	0.060 (0.039)	−0.082 (0.046)	INQ

(continued)

Table 3.1 (continued)

Country	Direct effects			Indirect effect	Model
	β_1 (SE)	β_2 (SE)	β_3 (SE)	$\beta_1 \times \beta_2$ (SE)	
United States of America	0.076 (0.063)	0.758 (0.029)*	−0.030 (0.050)	0.057 (0.048)	INQ
Ninth grade participants					
Botswana	−0.191 (0.127)	0.763 (0.067)*	0.156 (0.113)	−0.146 (0.098)	INQ
Honduras	0.295 (0.112)*	0.958 (0.033)*	−0.005 (0.065)	0.283 (0.110)*	MED
South Africa	−0.398 (0.067)*	0.818 (0.039)*	0.012 (0.061)	−0.325 (0.058)*	MED(−)
Benchmarking participants					
Abu Dhabi, UAE	−0.086 (0.110)	0.861 (0.033)*	0.095 (0.064)	−0.074 (0.095)	INQ
Alberta, Canada	0.160 (0.088)	0.801 (0.038)*	−0.033 (0.072)	0.128 (0.073)	INQ
Dubai, UAE	−0.136 (0.078)	0.814 (0.043)*	0.006 (0.064)	−0.111 (0.064)	INQ
Ontario, Canada	0.111 (0.117)	0.590 (0.070)*	0.158 (0.097)	0.065 (0.067)	INQ
Quebec, Canada	0.097 (0.081)	0.871 (0.029)*	0.101 (0.051)	0.084 (0.070)	PSC&INQ

Note (−) indicates a negative regression coefficient. *SE* standard error. *UAE* United Arab Emirates. *$p < 0.05$

both SEAS and order for all motivational aspects and model INQ was consistently found for the safety component. Nevertheless, there are countries where the models differ (consider Singapore and Kazakhstan).

Second, for SEAS and safety in schools, model INQ dominated; the MED model was less common, but could be identified for more than 20 % of the countries. In addition, model PSC&INQ was supported in 13 countries for students' self-concept and in 10 countries for intrinsic value. For order in schools, the MED model was the most common, but the INQ model was also apparent for a significant number of countries (Table 3.3).

Third, differences in the assignment of a specific model across the different aspects of achievement motivation could be identified. For instance, while the PSC&INQ model can be found in more than 20 % of the countries for SEAS and for the two motivational constructs of self-concept and intrinsic value, this model is substantially less common for students' extrinsic value. For the remaining combinations of scenarios, the overall frequencies across the motivational aspects are relatively consistent.

Fourth, we found support for model PSC in only two cases, which implies that only teachers' perceptions of the school climate were significantly related to motivation (Table 3.3).

Table 3.2 Overview of existing models describing the relations among school climate, instructional quality, and achievement motivation (see Fig. 3.2 for explanation of scenarios)

Countries	SEAS			Safety			Order		
	Self-concept	Intrinsic value	Extrinsic value	Self-concept	Intrinsic value	Extrinsic value	Self-concept	Intrinsic value	Extrinsic value
Armenia	N/A	INQ	MED(−)	N/A	MED(−)	MED(−)	N/A	MED	INQ
Australia	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Bahrain	INQ	PSC&INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Chile	PSC&INQ	INQ	INQ	PSC&INQ	INQ	INQ	INQ	INQ	INQ
Chinese Taipei	PSC&INQ	PSC&INQ	INQ	PSC&INQ	PSC&INQ	INQ	PSC&INQ	INQ	INQ
England	INQ	INQ	PSC&INQ	INQ	INQ	INQ	MED	MED	MED
Finland	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Georgia	PSC&INQ	PSC&INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Ghana	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Hong Kong SAR	MED	MED	MED	INQ	INQ	PSC&INQ	MED	MED	MED
Hungary	PSC&INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Indonesia	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Iran, Islamic Rep. Of	PSC&INQ	PSC&INQ	INQ	MED(−)	PSC&INQ	MED(−)	INQ	PSC&INQ	INQ
Israel	PSC&INQ	PSC&INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Italy	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Japan	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Jordan	INQ	INQ	INQ	INQ	MED(−)	MED(−)	INQ	INQ	INQ
Kazakhstan	PSC&INQ	PSC&INQ	INQ	INQ	PSC(−)&INQ	PSC(−)&INQ	MED	MED	INQ

(continued)

Table 3.2 (continued)

Countries	SEAS			Safety			Order		
Korea, Rep. Of	INQ	INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Lebanon	INQ	PSC&INQ	INQ	INQ	INQ	INQ	PSC&INQ	MED	INQ
Lithuania	INQ	INQ	INQ	MED	MED	MED	INQ	INQ	INQ
Macedonia	INQ	INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Malaysia	INQ	PSC&INQ	PSC&INQ	INQ	INQ	INQ	INQ	INQ	INQ
Morocco	PSC&INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
New Zealand	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Norway	MED	MED	MED	INQ	INQ	MED	MED	MED	MED
Oman	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Palestinian Nat'l Auth.	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Qatar	INQ	INQ	MED	INQ	INQ	INQ	MED	MED	MED
Romania	PSC&INQ	INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Russian Federation	MED	MED	MED	MED	MED	MED	MED	MED	MED
Saudi Arabia	INQ	INQ	INQ	INQ	INQ	PSC&INQ	PSC&INQ	INQ	INQ
Singapore	PSC&INQ	INQ	PSC(−)&INQ	PSC&INQ	INQ	INQ	PSC&INQ	INQ	INQ
Slovenia	PSC&INQ	PSC&INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Sweden	MED	MED	MED	INQ	INQ	INQ	MED	MED	MED
Syrian Arab Republic	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ

(continued)

Table 3.2 (continued)

Countries	SEAS			Safety			Order		
Thailand	INQ	INQ	INQ	PSC&INQ	INQ	PSC&INQ	MED	MED	MED
Tunisia	INQ	INQ	INQ	INQ	INQ	INQ	MED	MED	MED
Turkey	INQ	PSC(−)&INQ	INQ	INQ	INQ	PSC(−)&INQ	MED	MED	MED
Ukraine	MED	MED	MED	INQ	PSC(−)&INQ	INQ	INQ	INQ	INQ
United Arab Emirates	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
United States of America	PSC&INQ	INQ	INQ	INQ	PSC(−)&INQ	INQ	MED	MED	MED
Ninth grade participants									
Botswana	INQ	INQ	PSC	INQ	INQ	N/A	INQ	INQ	N/A
Honduras	MED	MED	MED	MED	MED	MED	MED	MED	MED
South Africa	MED(−)	MED(−)	MED(−)	MED(−)	MED(−)	MED(−)	PSC&INQ	PSC&INQ	INQ
Benchmarking participants									
Abu Dhabi, UAE	INQ	INQ	INQ	INQ	INQ	INQ	INQ	PSC&INQ	PSC&INQ
Alberta, Canada	INQ	INQ	INQ	INQ	PSC(−)&INQ	INQ	INQ	INQ	INQ
Dubai, UAE	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ	INQ
Ontario, Canada	PSC	INQ	PSC&INQ	INQ	INQ	INQ	N/A	MED	MED
Quebec, Canada	PSC&INQ	PSC&INQ	PSC&INQ	INQ	INQ	INQ	MED	MED	MED

Note (−) indicates a negative regression coefficient. *SE* standard error. *UAE* United Arab Emirates. *p < 0.05

Table 3.3 Absolute frequencies of the scenarios (see Fig. 3.2) identified in the 50 participating TIMSS 2011 countries

Model	SEAS			Safety			Order		
	Self-concept	Intrinsic value	Extrinsic value	Self-concept	Intrinsic value	Extrinsic value	Self-concept	Intrinsic value	Extrinsic value
INQ	22	26	29	40	38	36	19	20	24
MED	12	12	13	3	3	4	24	26	24
MED(−)	1	1	2	2	3	4	0	0	0
PSC	1	0	1	0	0	0	0	0	0
PSC&INQ	13	10	4	4	2	3	5	4	1
PSC(−)&INQ	0	1	1	0	4	2	0	0	0
N/A	1	0	0	1	0	1	2	0	1

Note (−) indicates a negative regression coefficient. *UAE* United Arab Emirates. *N/A* indicates that either the relations among the three constructs were insignificant ($p > 0.05$) or there was only a significant relation between school climate and instructional quality without any connection to the motivational outcome variable

Fifth, looking at the role of instructional quality as a mediator between aspects of school climate and achievement motivation, the following cultural patterns could be identified:

- *Scandinavian countries*: Mediation was apparent for SEAS and order across all motivational aspects.
- *English-speaking countries*: Mediation was apparent for SEAS and order across all motivational aspects in Australia and New Zealand. In England, the USA, and Quebec, only the relation between order and motivation was mediated by instructional quality.
- *Asian countries*: In Japan and Hong Kong, the relations between SEAS and achievement motivation, and order and achievement motivation were mediated by instructional quality; in Korea and Thailand, mediation was apparent only for order.
- *Eastern and Central European countries*: instructional quality mediated the relation between order and student motivation in Georgia, Romania, Macedonia, Slovenia, and Kazakhstan. In the Russian Federation and the Ukraine, SEAS was mediated; in the Russian Federation and Lithuania, safety in schools was mediated.
- *Arabic countries*: The relation between safety and student motivation was mediated in Iran and Jordan; in addition, SEAS and order were mediated in Oman, Palestine, and Qatar.
- *North Africa*: The relation between order and student motivation was mediated by instructional quality in Tunisia and Turkey; mediation was also found for SEAS and safety in South Africa.
- *South America*: The models with SEAS, safety, and order as school climate aspects showed mediation in Honduras.

Overall, the findings indicate that differing scenarios of relations exist among school climate, instructional quality, and achievement motivation. Although there were different patterns of relations across the aspects of school climate and achievement motivation, models INQ and MED dominate.

3.7 Discussion

This study was concerned with the relations among school climate, instructional quality, and achievement motivation across the 50 participating TIMSS 2011 countries in grade eight in mathematics. We proposed a research model that allowed us to identify four potential scenarios that indicated different substantive interpretations. With the help of MSEM, we found that models INQ and MED dominated across all aspects of school climate and achievement motivation.

As a major aim in practice, policy, and teacher education is to increase the level of instructional quality, as well as boost students' motivation for mathematics, our

findings emphasize the importance of school climate for instructional quality and motivation (Creemers and Kyriakides 2008; Sammons 2009). Collective beliefs, capabilities, and trust among the included members of the school institution, as manifested by high levels of SEAS, are important for instructional quality and achievement motivation. Creating a school climate that is oriented toward academic success can therefore be associated with higher instructional quality, which in turn leads to positive student outcomes. Although we do not claim causality in this mechanism, we believe that a positive school climate is indeed beneficial for instruction (Mitchell et al. 2010; Morin et al. 2014). In this respect, we found support for the mediating role of instructional quality in the relation between SEAS and achievement motivation in a number of countries. Nevertheless, in some cases, higher levels of SEAS were associated with lower instructional quality; this finding may point to the potential negative consequences of emphasizing academic success in such a way that competition and an orientation toward performance rather than motivation emerge.

Another explanation for the mediation may refer to the conceptualization and measurement of SEAS as an aspect of teachers' perceived school climate. SEAS was measured as a latent variable, where indicators refer to parents', students', and teachers' ambitions and priorities for learning and academic success. The covariance of these indicators reflects the collective and shared beliefs among these members of the school institution (Hoy et al. 2006). The link between these members that may arise when everyone aims for the same goal seems to influence teachers' instructional quality and student motivation. Indeed, previous research has shown that a strong student-teacher relationship positively influences student achievement (Roorda et al. 2011).

Moreover, we note that the mediation model was particularly apparent for the order component; this again indicates that order in schools may serve as a prerequisite for creating learning environments of high quality. But since the strength of the mediation differed across countries, further investigation is needed to assess what determines this mechanism in the context of the specific countries. Our secondary data analysis showed cross-country differences in the occurrence of mediation.

In summary, as previous research has found that instructional quality and school climate are related to both achievement student motivation (Fauth et al. 2014; Hoy et al. 2006; Klieme et al. 2009; Klusmann et al. 2008; Nilsen and Gustafsson 2014), it is thus unsurprising that we found that school climate influences motivation.

Although the importance of school climate for students' achievement motivation has been confirmed in our study, the effects of instructional quality dominated across almost all scenarios. Indeed, previous research has identified significant interactions between instructional quality and learning outcomes (Baumert et al. 2010; Blömeke et al. 2013; Fauth et al. 2014), and our results support these findings. Moreover, while the well-recognized research in this field is often restricted to German-speaking countries (see Baumert et al. 2010; Klieme et al. 2009) and English-speaking countries (Brophy 2006; Good et al. 2009), our findings support previous research in general, but also contribute to fostering understanding of

educational policy and practice in other countries as well, including developing countries.

The strong relation between instructional quality and achievement motivation may have two potential sources. First, instruction that focuses on engaging students to learn mathematics also creates opportunities for students to become motivated by the subject. Second, the measurement of instructional quality focused mainly on the engagement part of the construct, thus aligning with the measurement of achievement motivation; this alignment of the measures may have created similarities in how students understand and rate the items presented in the TIMSS 2011 questionnaire. Regardless, this strong association was consistently found in almost all countries and therefore needs further attention.

While there have been a number of studies on the relations between school climate and achievement (see Hoy et al. 2006; Martin et al. 2013; McGuigan and Hoy 2006) and on relations between instructional quality and achievement (e.g. Baumert et al. 2010), there have been relatively few studies investigating the relations between school climate and instructional quality (Creemers and Kyriakides 2010). The findings support our expectation of an association between instructional quality and school climate, and point out the importance of SEAS, Safety, and Order as important aspects of school climate. Moreover, including all countries and using international large-scale studies may also inform policy and practice about the importance of SEAS for instructional quality.

It is noteworthy that SEAS and instructional quality play an important role not only for students' motivation in learning mathematics but also for their self-beliefs and future-oriented motivation to pursue a career in mathematics and value the subject; this points to the significance of both the school and the classroom environment for career aspirations and for students' evaluation of their own capabilities in mathematics, which in turn determine their performance.

3.8   Limitations

One limitation of the present study is the measurement of the core construct, instructional quality. Although the student ratings of teachers' clarity and support in learning provide valid indicators of instructional quality with respect to instructional practices that are aimed at engaging students in learning mathematics (Scherer and Gustafsson 2015a), it is desirable to capture further aspects, such as cognitive activation and classroom management (Fauth et al. 2014; Klieme et al. 2009). We believe that gaining conceptual breadth in the measurement of instructional quality provides (a) a better representation of this multidimensional construct, and (b) more information about whether or not different aspects of instructional quality relate differently to student outcomes.

3.9　Conclusion

Besides supporting the importance of classroom instruction for motivational outcomes, our study advocates the relevance of school climate for both instructional quality and achievement motivation in many countries, feeding into the search for ways to improve instructional quality. We encourage further research into the field of educational effectiveness, to study the effects of instructional quality on educational outcomes by accounting for the school climate context.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.

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

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling: A Multidisciplinary Journal, 13*(2), 186–203. doi:10.1207/s15328007sem1302_2

Blömeke, S., Gustafsson, J. E., & Shavelson, R. (2013). Assessment of competencies in higher education. *Zeitschrift für Psychologie, 221*(3), 202.

Bofah, E. A.-t., & Hannula, M. S. (2015). TIMSS data in an African comparative perspective: Investigating the factors influencing achievement in mathematics and their psychometric properties. *Large-scale Assessments in Education, 3*(1), doi:10.1186/s40536-015-0014-y

Brophy, J. (2006). Observational research on generic aspects of classroom teaching. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 755–780). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. doi:10.1080/10705510701301834

Chen, P., & Vazsonyi, A. T. (2013). Future orientation, school contexts, and problem behaviors: A multilevel study. *Journal of Youth and Adolescence, 42*(1), 67–81. doi:10.1007/s10964-012-9785-4

Creemers, B., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: a contribution to policy, practice and theory in contemporary schools*. Abingdon, Oxon: Routledge.

Creemers, B., & Kyriakides, L. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness and School Improvement, 21*(4), 409–427.

Davison, M. L., & Srichantra, N. (1988). Acquiescence in components analysis and multidimensional scaling of self-rating items. *Applied Psychological Measurement, 12*(4), 339–351. doi:10.1177/014662168801200402

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*: Dordrecht: Springer Science & Business Media.

Deemer, S. (2004). Classroom goal orientation in high school classrooms: revealing links between teacher beliefs and classroom environments. *Educational Research, 46*(1), 73–90. doi:10.1080/0013188042000178836

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132.

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

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.

Goddard, R. D. (Ed.) (2002). *Collective efficacy and school organization: A multilevel analysis of teacher influence in schools* (Vol. 1). Greenwich: CT: Information Age Publishing.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., & Preckel, F. (2014). "My questionnaire is too long!" The assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*(3), 188–205. doi:10.1016/j.cedpsych.2014.04.002

Goldstein, S. E., Young, A., & Boyd, C. (2008). Relational aggression at school: Associations with school safety and social climate. *Journal of Youth and Adolescence, 37*(6), 641–654.

Good, T. L., Wiley, C. R., & Florez, I. R. (2009). Effective teaching: An emerging synthesis. In L. J. Saha & A. G. Dworkin (Eds.), *International handbook of research on teachers and teaching* (pp. 803–816). Dordrecht: Springer.

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: do they matter? *Personality and Individual Differences, 35*(6), 1241–1254. doi:10.1016/S0191-8869(02)00331-8

Gregory, A., Cornell, D., & Fan, X. (2012). Teacher safety and authoritative school climate in high schools. *American Journal of Education, 118*(4), 401–425.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: motivational and informational components. *Developmental Psychology, 17*(3), 300.

Hattie, J. (2009). *Visible learning: A synthesis of 800 + meta-analyses on achievement*. Abingdon, Oxon: Routledge.

Hoy, W. K., & Tschannen-Moran, M. (1999). Five faces of trust: An empirical confirmation in urban elementary schools. *Journal of School Leadership, 9*, 184–208.

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425–446.

IEA. (2012). *International database analyzer (version 3.1)*. (Software) Hamburg, Germany: International Association for the Evaluation of Educational Achievement (IEA). Retrieved from http://www.iea.nl/data.html.

Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: Perspectives from technology, economy, and educational research* (pp. 115–147). Dordrecht: Springer, Netherlands.

Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). New York, NY: Waxmann Publishing Co.

Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Teachers' occupational well-being and quality of instruction: The important role of self-regulatory

patterns. *Journal of Educational Psychology, 100*(3), 702–715. doi:10.1037/0022-0663.100.3.702

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology, 105*(3), 805–820. doi:10.1037/a0032583

Kythreotis, A., Pashiardis, P., & Kyriakides, L. (2010). The influence of school leadership styles and culture on students' achievement in Cyprus primary schools. *Journal of Educational Administration, 48*(2), 218–240.

Lazarides, R., & Ittel, A. (2012). Instructional quality and attitudes toward mathematics: Do self-concept and interest differ across students' patterns of perceived instructional quality in mathematics classrooms? *Child Development Research, 2012*, 1–11. doi:10.1155/2012/813920

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 modeling. *Contemporary Educational Psychology, 34*(2), 120–131. doi:10.1016/j.cedpsych.2008.12.001

Marsh, H. W., & Gouvernet, P. J. (1989). Multidimensional self-concepts and perceptions of control: Construct validation of responses by children. *Journal of Educational Psychology, 81*(1), 57–69.

Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F., Leung, K. C., & Parker, P. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*(1), 108–128. doi:10.1037/a0029907

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist, 47*(2), 106–124. doi:10.1080/00461520.2012.670488

Martin, M. O., Foy, P., Mullis, I. V. S., & O'Dwyer, L. M. (2013). Effective schools in reading, mathematics, and science at fourth grade. In M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning* (pp. 109–178). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martin, M.O. & Mullis, I.V.S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

McGuigan, L., & Hoy, W. K. (2006). Principal leadership: Creating a culture of academic optimism to improve achievement for all students. *Leadership and Policy in Schools, 5*(3), 203–229. doi:10.1080/15700760600805816

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006a). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487–503. doi:10.1146/annurev.psych.56.091103.070258

Meece, J. L., Glienke, B. B., & Burg, S. (2006b). Gender and motivation. *Journal of School Psychology, 44*(5), 351–373. doi:10.1016/j.jsp.2006.04.004

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Mitchell, M. M., & Bradshaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: The role of classroom management and exclusionary discipline strategies. *Journal of School Psychology, 51*(5), 599–610. doi:10.1016/j.jsp.2013.05.005

Mitchell, M. M., Bradshaw, C. P., & Leaf, P. J. (2010). Student and teacher perceptions of school climate: A multilevel exploration of patterns of discrepancy. *Journal of School Health, 80*(6), 271–279. doi:10.1111/j.1746-1561.2010.00501.x

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–24. doi:10.1080/10705511.2014.961800

Morin, A. J. S., Marsh, H. W., Nagengast, B., & Scalas, L. F. (2014). Doubly latent multilevel analyses of classroom climate: An illustration. *The Journal of Experimental Education, 82*(2), 143–167. doi:10.1080/00220973.2013.769412

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muthén, B., & Muthén, L. (1998–2014). *Mplus Version 7.3*. Los Angeles, CA: Muthén & Muthén.

Nilsen, T., & Gustafsson, J.-E. (2014). School emphasis on academic success: exploring changes in science performance in Norway between 2007 and 2011 employing two-level SEM. *Educational Research and Evaluation, 20*(4), 308–327. doi:10.1080/13803611.2014.941371

NSF. (2012). *Science and engineering indicators 2012*. Retrieved 31 December 2015 from http://www.nsf.gov/statistics/seind12/.

OECD. (2007). *PISA 2006. Science competencies for tomorrow's world* (Vol. 1). Paris: OECD Publishing.

OECD. (2014a). *Education at a glance 2014: OECD indicators*. Paris: OECD Publishing. doi:10.1787/eag-2014-en

OECD. (2014b). *TALIS 2013 results*. Paris: OECD Publishing.

Pintrich, P., & Schunk, D. (2002). *Motivation in education: Theory. research, and applications* (2nd ed.). Columbus, OH: Merrill Prentice Hall.

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. doi:10.1037/0021-9010.88.5.879

Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods, 15*(3), 209–233. doi:10.1037/a0020141

Preckel, F. (2014). Assessing need for cognition in early adolescence. *European Journal of Psychological Assessment, 30*(1), 65–72. doi:10.1027/1015-5759/a000170

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354–373. doi:10.1037/a0029315

Roorda, D. L., Koomen, H. M., Spilt, J. L., & Oort, F. J. (2011). The influence of affective teacher–student relationships on students' school engagement and achievement a meta-analytic approach. *Review of Educational Research, 81*(4), 493–529.

Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement, 74*(1), 31–57. doi:10.1177/0013164413498257

Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00081

Sammons, P. (2009). The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools. *School Improvement and School Effectiveness, 20*(1), 123–129.

Scherer, R., & Gustafsson, J.-E. (2015a). Student assessment of teaching as a source of information about aspects of teaching quality in multiple subject domains: An application of multilevel bifactor structural equation modeling. *Frontiers in Psychology, 6*(1550). doi:10.3389/fpsyg.2015.01550

Scherer, R., & Gustafsson, J.-E. (2015b). The relations among openness, perseverance, and performance in creative problem solving: A substantive-methodological approach. *Thinking Skills and Creativity, 18*, 4–17. doi:10.1016/j.tsc.2015.04.004

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. doi:10.3102/0034654307310317

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*(1), 70–83. doi:10.1037/0012-1649.42.1.70

Stancel-Piątak, A., & Desa, D. (2014). Methodological implementation of multi group multilevel SEM with PIRLS 2011: Improving reading achievement. In R. Strietholt, W. Bos, J.-E. Gustafsson, & M. Rosén (Eds.), *Educational policy evaluation through international comparative assessments* (pp. 75–91). Muenster, New York: Waxmann.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research, 83*(3), 357–385.

van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. *PLoS ONE, 8*(7), 1–7. doi:10.1371/journal.pone.0068967

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2015). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology,*. doi:10.1037/edu0000075

Wang, M.-T., & Degol, J. L. (2015). School climate: a review of the construct, measurement, and impact on student outcomes. *Educational Psychology Review,* 1–38. doi:10.1007/s10648-015-9319-1

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. doi:10.1006/ceps.1999.1015

Wigfield, A., Battle, A., Keller, L. B., & Eccles, J. (2002). Sex differences in motivation, self-concept, career aspirations and career choice: Implications for cognitive development. In A. V. McGillicuddy-De Lisi & R. De-Lisi (Eds.), *Biology, sociology, and behavior: The development of sex differences in cognition* (pp. 93–124). Greenwich, CT: Ablex.

Wilson, D. (2004). The interface of school climate and school connectedness and relationships with aggression and victimization. *Journal of School Health, 74*(7), 293–299.