

Facilitating Interest Development and Achievement in Classroom Instruction: How Large-Scale
Studies Like TIMSS Can be Used to Investigate Effects of Instructional Quality

Mareike Kunter & Yi-Miau Tsai

Max Planck Institute for Human Development, Berlin, Germany

Facilitating Interest Development and Achievement in Classroom Instruction: How Large-Scale Studies Like TIMSS Can be Used to Investigate Effects of Instructional Quality

Abstract

This study uses data from the 1995 TIMSS middle school study to investigate efficient use of classroom time and personal support of students as important aspects of instructional quality. Typically, most studies on instructional quality define quality in terms of instructional effects either on cognitive outcomes, such as achievement gains, or on non-cognitive outcomes, such as students' motivational development, hypothesizing that different aspects of instruction show specific effects on each set of variables. In the present study, we examine whether mathematical instruction that is characterized by efficient use of time but, at the same, leaves room for personal support fosters both mathematical achievement and mathematics-related interest. In addition, we investigate the hypothesis that an imbalance between the two instructional aspects may have specific unfavourable effects on the two outcomes. The study draws on the German TIMSS sample of 8th graders and a national longitudinal extension which included 1900 students for whom additional data from a measurement point in grade 7 and supplementary questionnaires were available. Students' class mean ratings of classroom instruction were used to predict learning gains and interest development over the course of the year. Results showed that efficient time use and teacher personal support affected students' achievement and interest differentially. Students in classrooms with less efficient time use showed weaker achievement gains, whereas personal support was found to have a positive effect on interest development. However, no empirical support was found for the expected interaction effects between the two aspects of instruction. These distinct effects highlight the need for multiple criteria of instructional quality. Considering motivational and achievement outcomes at the same time may usefully inform educators about good teaching.

Keywords: interest, student achievement, instructional quality

Introduction

In instructional research, students' achievement and motivation are often investigated separately. *Achievement and learning gains* have been the focus of school effectiveness studies and studies on instructional quality, which have shown that schools and teachers differ in the degree to which they foster student learning (Anderson, Ryan, & Shapiro, 1989; Luyten, Visscher, & Witziers, 2005; Wang, Haertel, & Walberg, 1993). As a core dimension of instructional quality, the provision of sufficient learning time has been shown to be one of the most important predictors for student achievement gains (Anderson et al., 1989; Seidel & Shavelson, 2007; Wang et al., 1993). It has been shown that teachers use a range of classroom management strategies to maximize time on task by preventing disciplinary problems and other potential time losses—for example, establishing clear rules, monitoring student behaviour, and reacting promptly to student misbehaviour (Doyle, 1986; Emmer & Stough, 2001).

Whereas the beneficial effects of effective classroom management and efficient time use on students' achievement are commonly acknowledged, their meaning for *students' motivation* is less clear. Motivational research, which typically draws on smaller-scale studies or experimental approaches, has focused on educational outcomes such as domain-specific interest and its sensitivity to instruction (Pintrich, 2003). For instance, individual interest as a dispositional preference for certain topics has been linked to higher engagement and persistence, which may in the long run support learning (Hidi, 2006; Krapp, 2002a). One widespread theoretical concern regarding the relation between efficient time use and interest relates to the high degree of control and external regulation that seem to be inherent in most classroom management approaches (Elias & Schwab, 2006; McCaslin et al., 2006). An emphasis on maximizing time on task through teachers' controlling behaviour, which in turn underlines the asymmetrical relationship between teachers and students, would seem to preclude the emergence of caring and supporting teacher–student relations. Such relations have been shown to boost students' development in non-achievement related outcomes such as intrinsic motivation, interest and well-being (Anderman & Kaplan, 2008; Davis, 2003; Furrer & Skinner, 2003). It has been argued that the teacher- and instruction-, but not student-centered style (Elias & Schwab, 2006) that seems typical for efficient classroom management may appear controlling to students, undermine their sense of autonomy, and thus thwart their intrinsic motivation or interest (Davis, 2003; Elias & Schwab, 2006; Reeve, 2006).

From a multi-criterial perspective, classroom instruction in which teachers emphasize efficient time use thus seems to be a double-edged sword. In terms of achievement, it can be expected that students in classes where much time is devoted to task work can achieve high

learning rates and achievement gains. In terms of students' motivation and, in particular, aspects of intrinsic motivation such as domain-specific interest, however, a strong focus on teacher regulation and management of student interactions, which is often a prerequisite for efficient time use, may be problematic—especially if it is perceived as controlling by the students. The key question that arises is thus not whether time is used efficiently or not, but whether this maximizing of task time is achieved at the expense of positive student–teacher relationships, leaving students feeling controlled rather than supported. Several empirical studies indicate that teachers' regulatory behaviour does not necessarily undermine students' motivation, but that if teachers respect students' autonomy and enforce classroom management strategies in a caring and supportive way, this instructional style can boost students' satisfaction, well-being and motivation (Kunter, Baumert, & Köller, 2007; Reeve, Bolt, & Cai, 1999; Skinner & Belmont, 1993; Walker, 2008; Franz E. Weinert & Andreas Helmke, 1995). For instance, in a small-scale study using observational data, teacher interviews and student ratings, Walker (2008) identified different types of teacher practices and found that students who showed the highest achievement, engagement and self-efficacy were taught by a teacher whose teaching style was characterised by consistent classroom management coupled with high levels of student support. Similarly, Weinert and Helmke (1995) employed observer ratings to describe teachers' regulatory behaviour and found that “supportive” teacher control, involving detailed instructions and frequent monitoring, was associated with positive attitudes towards the subject and active task engagement. In addition, there is evidence that providing students with structure through clear and consistent classroom management is positively linked with a sense of autonomy and competence (Kunter et al., 2007; Skinner & Belmont, 1993). These studies thus underscore that a supportive approach to enforcing order and discipline in class may have positive effects on students' motivation as well as their achievement. However, as most of these studies were based on cross-sectional data with small selective samples of teachers and students and rarely assessed achievement and motivation measures simultaneously, studies with higher external validity are needed to support this claim.

In this paper, we therefore investigate efficient time use and personal support of students as features of classroom instruction in terms of their potential to foster both student achievement and subject-related interest. The 1995 TIMSS study in Germany, with its representative class sample, its standardized achievement tests and additional student questionnaires, provides a unique data base to investigate our hypotheses. Moreover, the additional longitudinal component integrated in the German TIMSS design is suitable for investigating the long-term impact of instructional features on educational outcomes while controlling for prior levels of achievement and motivation.

The study is an extension of prior studies on instructional quality using the German TIMSS data in which selected instructional aspects were investigated separately (Kunter & Baumert, 2006; Kunter et al., 2007). Based on these results, we first investigate the relationships among efficient time use and personal support of students to determine whether the two aspects coincide or are mutually exclusive. In so doing, we address the assumption that efficient time use is most often achieved by intrusive regulation of student behaviour and a strong focus on academic matters that leaves little room for personal support (Kohn, 1996; McCaslin & Good, 1996). Second, we examine the long-term effect of both aspects on students' achievement and interest development. Drawing on research that documents the positive effect of efficient time use on student achievement, on the one hand (Anderson et al., 1989; Seidel & Shavelson, 2007; Wang et al., 1993), and the beneficial effects of positive teacher–student relations on various aspects of students' motivation, on the other being (Davis, 2003; Furrer & Skinner, 2003), we expect to find that mathematical instruction characterized by both efficient use of time and a high degree of personal support for students, even in non-academic matters, will have positive effects on students' achievement as well as on their mathematics-related interest. Further, we extend prior research by specifically investigating the interaction between the two aspects of instruction. Specifically, we expect students in classes where efficient time use is emphasized, but teachers do not invest time in supporting students in personal matters, to show lower interest. We also expect students in classes in which teachers strive to support their students in personal matters, but do not succeed in using time efficiently, to have particularly low achievement gains.

Methodology and Data Sources

Data Set and Participants

This study draws on data from the German sample of the TIMSS middle school assessment (Beaton et al., 1996). In Germany, the international cross-sectional design of the TIMSS study was extended to a longitudinal design, with students being first assessed in grade 7 and again one year later as part of the international TIMSS assessment in grade 8. The longitudinal sample comprises 1900 students (48.5% girls) from 80 classrooms in 80 schools of all academic tracks. The students' mean age at the second measurement point (grade 8) was 14.8 years. The mean number of students per class was 24 (range: 13-31). Overall, 34% of the students were enrolled in vocational-track schools (*Hauptschule*), 28% in intermediate-track schools (*Realschule*), and 39% in the academic track (*Gymnasium*). Relative to the nationally representative cross-sectional TIMSS sample, this sample is slightly biased towards the academic track.

Measures

Mathematics achievement was assessed with the TIMSS achievement test, which covers content areas and performance categories such as conducting routine and complex procedures, applying knowledge and solving mathematical problems (Beaton et al., 1996). Students were administered overlapping item in grade 7 and grade 8. At both times, items from the international item pool were combined with a set of items taken from the First and the Second International Mathematics Study (Husén, 1967; Robitaille & Garden, 1989) and an earlier German study by the Max Planck Institute for Human Development. Individual achievement scores were estimated based on item response theory (see Martin & Kelly, 1997, for details). To describe students' achievement on a single dimension, we used an equating procedure in which achievement at grade 8 served as a scale anchor.

Mathematics-specific interest was assessed in grade 7 and grade 8 using a four-item scale tapping aspects of value commitment and positive emotion in line with the person-object approach to interest (Krapp, 2002b). Students answered questions on positive feelings and perceived importance of mathematics on a 5-point scale. The internal consistency of the scale was .81 in grade 7 and .78 in grade 8 (Cronbach's alpha).

Two scales from a student questionnaire were used to gain measures of instructional quality in each of the classes. Items on efficient time use tapped the amount of time lost through disciplinary problems and other disturbances in class (e.g., "At the beginning of mathematics class, it takes quite a while until students calm down and start working"; four items, all recoded, Cronbach's alpha = .72). Items on personal support tapped teachers' willingness to provide help and support unrelated to content matters (e.g., "Our teacher soon notices if a student has problems or worries"; 3 items, Cronbach's alpha = .87). Students' ratings from each class were averaged to provide an objective measure of instruction. We drew on three indicators to establish the feasibility of using student ratings as indicators for the quality of classroom instruction (Lüdtke, Trautwein, Kunter, & Baumert, 2006 see Table 1). First, the intraclass correlations (ICC_1) indicated a substantial degree of between-class variation for both aspects of instruction. Second, the reliability of the class means as indicated by the ICC_2 (applying the Spearman Brown Formula to the ICC_1 by taking into account an average of 24 ratings per class) was sufficient (as a rule of thumb, an $ICC_2 > .70$ is considered to indicate good reliability). Third, the average deviation index AD_M showed that, on average, students deviated by two thirds of a scale point in their ratings within classes, which can be considered to indicate substantial agreement. In sum, these indicators demonstrate high inter-rater agreement for the students and confirm that it was appropriate to use mean class ratings as measures for instruction in class.

School track: Students from different tracks are likely to differ in their achievement levels, as well as in other variables. To account for these differences, school track was included as two dummy variables: the first (high track) coded 1 for classes in the academic track (*Gymnasium*) and 0 otherwise, and the second (low track) coded 1 for classes in the vocational track (*Hauptschule*) and 0 otherwise. Including these two dummy variables in the multilevel regression analyses allows us to describe the differences of those two tracks relative to the intermediate track (*Realschule*).

Statistical Approach

Multilevel Structure. The data have a multilevel structure, with students nested within classes. At Level 1, data on mathematics achievement and mathematics-related interest were obtained from each student. At Level 2, data on instructional features were obtained by averaging student responses on efficient time use and personal support across classrooms (school track was included in additional models). Multilevel modelling techniques were used to handle these multilevel data, which violate the principle of independence. All analyses were carried out using the Mplus software (Version 5.1 Muthén & Muthén, 1998-2004).

Missing Data. Due to the longitudinal design and the combination of performance measures and questionnaire data, complete data sets were not available for all students (see Table 1). A small proportion of missing data was caused by the drop-out of low-achieving students at the second measurement point, but most of the missing data seemed to occur at random (Kunter, 2005). In order to avoid a reduction of the sample size, missing values were imputed and complete datasets were produced for all students (Rubin, 1996). Multiple imputation was performed using the NORM software (Schafer, 1999). Specifically, missing values were imputed using a number of auxiliary variables (student background variables, mathematics achievement score, additional motivational and classroom perception scales) and five complete data sets were produced for all students. Using the Mplus software, analyses were then carried out with these multiple datasets simultaneously and combined parameters were produced.

Findings

Descriptive statistics for student achievement and mathematics-related interest, as well as student ratings of time use and personal support are presented in Table 1. The between-class variability in mathematics achievement and interest as indicated by the ICC₁ shows that students differ substantially between classes, partly as an effect of the tracked system. As expected, student achievement differed according to track, with highest achievement in the *Gymnasium*

track, average achievement in *Realschule*, and lowest achievement in *Hauptschule* (see Table 2). *Gymnasium* students reported slightly lower mathematics-related interest than the other students, a phenomenon that has been repeatedly described for German students (see also Pekrun & Zirngibl, 2004). Patterns of instructional quality also differed between school tracks, with more efficient time use in academic track classes and higher personal support in vocational-track classes (Kunter & Baumert, 2006).

Regarding our first research question on the interplay between efficient time use and personal support, the school-type specific pattern points to the existence of specific teaching styles in which one aspect is stressed at the expense of the other. However, across the whole sample, we found a positive correlation of $r = .27$ ($p < .05$) between efficient time use and student support. Within tracks, we found positive correlations, although effect sizes varied and, with the exception of *Gymnasium* track, did not reach statistical significance ($r = .56$, $p < .05$ in *Gymnasium*; $r = .12$, $p > .05$ in *Realschule*; $r = .22$, $p > .05$ in *Hauptschule*). This result indicates that, overall, the creation of a classroom structure in which time is used efficiently does not necessarily mean costs for positive personal relationships with students. Instead, the positive correlation indicates that both aspects constitute related aspects of instructional quality that may coincide—or may both be absent. However, it seems that teachers in the specific tracks tend to stress either one of the two aspects in their teaching styles. Whereas this interaction between school type and instructional style is an interesting descriptive finding per se (Kunter et al., 2005, for a detailed discussion of differences in instructional quality across school tracks), it is also a limitation for the following regression analyses, as the high amount of confounded variance will reduce the power to detect unique effects for the instructional variables.

Regarding our second research question, multilevel path analyses were conducted to predict students' mathematics achievement and mathematics-related interest in grade 8. In these analyses, mathematics achievement and interest in grade 8 were modelled as two simultaneous outcome criteria on the individual and class level. On the individual level, achievement in grade 7 and interest in grade 7 were entered as control variables for achievement in grade 8 and interest in grade 8 respectively. As a base model to establish the between-class variance in students' achievement and interest changes over the course of the school year, we ran a model including only these individual level variables, which showed that 27% of the variance in achievement change and 4% of the variance in interest change was located at the between-class level (both $p < .05$). Although these effect sizes are small, they indicate that, even when prior achievement or interest are controlled, classes still differ systematically in the size of achievement or interest gains. In the next steps, we thus investigated whether changes in achievement and interest could

be predicted by instructional variables. On the class level, students' aggregated ratings for efficient time and personal support were entered as predictors. As an additional predictor, the interaction between time use and personal support (formed by the product of the two standardized predictors) was included. Table 2 shows that both instructional variables were associated with grade 8 achievement. Whereas efficient time use was positively associated with achievement one year later, a negative effect was found for personal support. These results indicate that students in classes reporting high time on task but low personal support show higher achievement gains over the course of one year than their peers in other classes. In contrast, only personal support had a positive effect on the motivational outcome of mathematics-related interest development ($b = .15$, $p < .01$). Contrary to our expectations, the interaction term did not emerge as a significant predictor of either achievement or interest.

As both student outcomes and instruction were found to differ between school tracks, in the next step, we included track as an additional class-level predictor. As shown in Model 2 in Table 2, all effects of instruction were reduced when school track was entered as a control variable. In particular, the effect for classroom management was no longer statistically significant in predicting student achievement. However, the negative coefficient for personal support in predicting student achievement remained significant.

Two approaches are helpful to evaluate the practical significance of the effects found (Kirk, 1996). First, the proportion of explained variance describes the explanatory value of the predictor sets at both levels. The explained variance on the class level is of particular interest here. Taking the between-class variance from the base model as a starting point, our first model explained 36% of the between-class variance in achievement and 37% of the between-class variance in interest. In a second base model, we estimated the variance explained by school track only, and found that 66% of the variance in achievement change and 12.5% of the variance in interest change could be attributed to track membership. In Model 2, which included both track and instruction, the model explained 71% of between-class variance in achievement and 49% of between-class variance in interest. Comparison with the second base model shows that the unique contribution of the instructional variables was 5% for achievement and 36% for interest, showing that, in general, instruction had moderate explanatory value over and above track membership, particularly on the outcome variable of mathematics-related interest.

As a second approach to evaluating the practical significance of each statistically significant predictor variable, we chose a procedure based on Tymms' (1997) proposal for calculating effect sizes for continuous level-2 predictors in multilevel models. This effect size, which is comparable with Cohen's d , can be calculated using the following formula:

$$\Delta = 2 * B * SD_{\text{predictor}} / \sigma_e$$

where B is the unstandardized regression coefficient in the multilevel model, $SD_{\text{predictor}}$ is the standard deviation of the predictor variable at the class level, and σ_e is the residual standard deviation at the student level. The resulting effect size describes the difference in the dependent variable between two classes that differ two standard deviations on the predictor variable. For efficient time use, we found an effect of $d = .54$ on achievement ($d = .15$ controlling for school track). For social support, we found an effect of $d = -.65$ on achievement ($d = -.29$ controlling for school track) and of $d = .33$ on interest ($d = .27$ controlling for school track). These can be classified as small to moderate effects.

Discussion: Educational and Scientific Importance

In this study, we investigated the specific effects of efficient use of classroom time and personal support of students on motivational development and learning in a large, representative sample of mathematics students in Germany. As a first finding, our results contradict the widespread assumption that efficient time use can only be achieved by sacrificing the dimension of personal student support. Instead, they suggest that teachers can orchestrate instructional activities to include various positive features, such as making efficient use of classroom time while at the same time providing individual students with guidance and support. Regarding our second research question on the long-term effects of the instructional features, our findings on the one hand replicated those of previous studies using individual perceptions and smaller samples, and specifically highlighted the distinct beneficial effect of interpersonal support for fostering students' subject-matter interest, and as a tendency also the positive effects of classroom management on achievement. On the other hand, we found an unexpected negative main effect of personal support on student achievement. To some degree, this effect can be explained by the track-specific pattern of low social support in the high-achieving *Gymnasium* classes. Still, even within tracks, a moderate negative effect was observed, which may indicate that a high emphasis on students' personal matters during class time may have costs for achievement. This finding was contrary to our expectations; we expected only a combination of inefficient time use and high support to be detrimental, but the interaction term between efficient time use and support was not found to be significant. There are two possible interpretations of this finding. On the one hand, it needs to be seen in the light of the limited time generally available in class. Note that we used aggregated student ratings as a measure of social support, meaning that the whole class shared the

impression that their teacher was willing to devote time to their non-academic problems. Given the high time pressures on teachers, who often complain that there is barely enough time to get through the required curriculum (e.g., Blase, 1986), it is possible that there simply is not enough time to attend to both academic and personal matters in class. The negative effect of social support as rated by the whole class in this study may indicate that attending to the whole class's personal problems may actually consume too much classroom time. However, this does not imply that teachers should not attend to students' personal matters at all—clearly personal issues can arise that teachers will need to deal with. The question to be addressed from the efficiency point of view is whether these issues need to be addressed immediately in the whole class context or would be better handled in individual teacher–student sessions.

A second interpretation also addresses a general limitation of our study. Given that we used a quasi-experimental design in which instruction was measured at only one time point (grade 8), we cannot rule out the possibility that the negative association between achievement gains and social support is a sign that the teachers are reacting to the students' growing academic difficulties. Longitudinal studies with several measurement points, and particularly including measurement at the beginning of the school year, are necessary to address this issue. As an additional limitation, issues of statistical power need to be discussed. High statistical power is needed to detect interaction effects and, in our study, the high occurrence of confounded variance (also a consequence of the quasi-experimental approach) and rather small sample size on the class level compromised statistical power. Again, further studies that disentangle effects and perhaps use a theory-based stratified sampling procedure to increase statistical power with respect to crucial theoretical variables would complement our results.

Despite these limitations, our study demonstrated how large-scale data can be productively used to investigate substantive theoretical questions in education. Our access to a representative sample of mathematics classes and use of the TIMSS achievement test as a highly objective and valid achievement measure strengthens the practical significance and generalizability of our findings. Thus, our findings complement existing studies on instructional quality, which are often based on small-scale studies with convenience samples. In addition, our findings on the differential effects of instruction may be particularly informative for the research community in the context of international large-scale comparisons. Although the interplay between learning and motivation is widely acknowledged (Eccles & Wigfield, 2002; Pintrich & Schunk, 1996; Stipek, 1996), students' domain-specific motivation and other psychosocial aspects representing desirable educational outcomes have rarely been investigated in large-scale studies (Gadeyne, Ghesquire, & Onghena, 2006). Our findings stress the importance of considering these additional

educational outcomes, over and above student achievement, to develop a full picture of what constitutes powerful education .

References

- Anderman, L. H., & Kaplan, A. (2008). The Role of Interpersonal Relationships in Student Motivation: Introduction to the Special Issue. *Journal of Experimental Education*, 76(2), 115-119.
- Anderson, L. W., Ryan, D., & Shapiro, B. (Eds.). (1989). *The IEA classroom environment study*. New York: Pergamon Press.
- Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- Blase, J. J. (1986). A qualitative analysis of sources of teacher stress: consequences for performance. *American Educational Research Journal*, 23(1), 13–40.
- Davis, H. A. (2003). Conceptualizing the Role and Influence of Student-Teacher Relationships on Children's Social and Cognitive Development. *Educational Psychologist*, 38(4), 207 - 234.
- Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching. A project of the American Educational Research Association* (3rd ed., pp. 392–431). New York: Macmillan.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109-132.
- Elias, M. J., & Schwab, Y. (2006). From Compliance to Responsibility: Social and Emotional Learning and Classroom Management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management* (pp. 309-341). Mahwah, N. J.: Lawrence Erlbaum.
- 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.
- Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95(1), 148-162.
- Gadeyne, E., Ghesquire, P., & Onghena, P. (2006). Psychosocial Educational Effectiveness Criteria and Their Relation to Teaching in Primary Education. *School Effectiveness and School Improvement*, 17(1), 63-85.
- Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review*, 1(2), 69–82.
- Husén, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries*. New York: John Wiley & Sons.
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56(5), 746-759.
- Kohn, A. (1996). *Beyond Discipline*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Krapp, A. (2002a). An educational-psychological theory of interest and its relation to SDT. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 405-427). Rochester, NY: University of Rochester Press.

- Krapp, A. (2002b). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12*(4), 383–409.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. Münster: Waxmann.
- Kunter, M., & Baumert, J. (2006). Linking TIMSS to research on learning and instruction: A reanalysis of the German TIMSS and TIMSS video data. In S. J. Howie & T. Plomp (Eds.), *Learning mathematics and science: Lessons learned from TIMSS*. (pp. 335-351). London: Routledge.
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective Classroom Management and the Development of Subject-Related Interest. *Learning & Instruction, 17*(5), 494-509.
- Kunter, M., Brunner, M., Baumert, J., Klusmann, U., Krauss, S., Blum, W., Jordan, A., & Neubrand, M. (2005). Der Mathematikunterricht der PISA-Schülerinnen und -Schüler: Schulformunterschiede in der Unterrichtsqualität. *Zeitschrift für Erziehungswissenschaft, 8*(4), 502–520.
- Lüdtke, O., Trautwein, U., Kunter, M., & Baumert, J. (2006). Reliability and agreement of student ratings of the classroom environment – A reanalysis of TIMSS data. *Learning Environments Research, 9*, 215-230.
- Luyten, H., Visscher, A., & Witziers, B. (2005). School Effectiveness Research: From a Review of the Criticism to Recommendations for Further Development. *School Effectiveness and School Improvement, 16*(3), 249-279.
- Martin, M. O., & Kelly, D. L. (Eds.). (1997). *Third International Mathematics and Science Study. Technical report: Vol II. Implementation and analysis. Primary and middle school years*. Chestnut Hill, MA: Boston College.
- McCaslin, M., Bozack, A. R., Napoleon, L., Thomas, A., Vasquez, V., Wayman, V., & Zhang, J. (2006). Self-Regulated Learning and Classroom Management: Theory, Research, and Considerations for Classroom Practice. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of Classroom Management* (pp. 223-252). Mahwah, N. J.: Lawrence Erlbaum.
- McCaslin, M., & Good, T. L. (1996). The informal curriculum. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 622-670). New York: Prentice Hall International.
- Muthén, L. K., & Muthén, B. O. (1998-2004). *Mplus user's guide* (3rd ed.). Los Angeles, CA.: Muthén & Muthén.
- Pekrun, R., & Zirngibl, A. (2004). Schülermerkmale im Fach Mathematik. In PISA-Konsortium Deutschland (Ed.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland -- Ergebnisse des zweiten internationalen Vergleichs* (pp. 191-210). Münster: Waxmann.
- Pintrich, P. R. (2003). Motivation and classroom learning. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of Psychology: Vol. 7. Educational Psychology* (pp. 103-122). Hoboken, NJ: Wiley.
- Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in Education*. Englewood Cliffs, NJ, USA: Prentice Hall.
- Reeve, J. (2006). Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit. *Elementary School Journal, 106*, 225-236.
- Reeve, J., Bolt, E., & Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology, 91*(3), 537-548.
- Robitaille, D. F., & Garden, R. A. (Eds.). (1989). *The IEA Study of Mathematics II: Contexts and outcomes of school mathematics*. Oxford: Pergamon Press.
- Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*(343), 473-489.
- Schafer, J. L. (1999). NORM for Windows 95/98/NT. Multiple imputation of incomplete data under a normal model. University Park, PA: Penn State Department of Statistics.

- 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.
- Skinner, E., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology, 85*(4), 571-581.
- Stipek, D. J. (1996). Motivation and Instruction. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 85–113). New York, NY: Macmillan.
- Tymms, P., Merrell, C., & Henderson, B. (1997). The First Year at School: A Quantitative Investigation of the Attainment and Progress of Pupils. *Educational Research and Evaluation, 3*(2), 101-118.
- Walker, J. M. T. (2008). Looking at Teacher Practices Through the Lens of Parenting Style. *Journal of Experimental Education, 76*(2), 218-240.
- 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.
- Weinert, F. E., & Helmke, A. (1995). Interclassroom differences in instructional quality and interindividual differences in cognitive development. *Educational Psychologist, 30*(1), 15–20.
- Weinert, F. E., & Helmke, A. (1995). Learning from wise Mother Nature or Big Brother Instructor: The wrong choice as seen from an educational perspective. *Educational Psychologist, 30*(3), 135–142.

Table 1: Descriptives, intercorrelations, and reliability information (class level) for the study variables

	N valid	M	SD	ICC ₁	Correlations		
					1.	2.	3.
<u>Individual variables</u>							
1. Achievement grade 7	1534	-0.17	1.08	0.43		-	-
2. Achievement grade 8	1798	0.45	1.20	0.52	0.94		-
3. Math-related interest grade 7	1527	3.23	0.83	0.10	-0.36	-0.42	1
4. Math-related interest grade 8	1383	3.16	0.87	0.10	-0.49	-0.53	0.80
<u>Class-level variables</u>							
				ICC ₁	ICC ₂	AD _M	
5. Efficient time use	80	2.43	0.31	0.130	0.78	0.64	
6. Personal support	80	2.47	0.34	0.171	0.83	0.67	

Note: ICC₁: Between-class variance; ICC₂: Reliability of class mean; AD_M: Average deviation index (agreement within classes); see text for further explanation.

Table 2: Multilevel models predicting students' mathematics achievement and interest in grade 8, standardized path coefficients B (standard errors of B)

Criterion	Model 1: Without School Track				M2: Controlling for School Track			
	Math achievement		Math-related interest		Math achievement		Math-related interest	
	B	SE	B	SE	B	SE	B	SE
<u>Individual variables</u>								
Achievement grade 7	0.512***	0.035	-	-	0.499***	0.034	0.539***	0.027
Math-related interest grade 7	-	-	0.539***	0.027				
<u>Class-level variables</u>								
Efficient time use	0.651***	0.170	0.032	0.090	0.182	0.126	0.119	0.095
Personal support	-0.711***	0.157	0.351***	0.070	-0.316*	0.128	0.283***	0.078
Interaction Efficient time use *								
Social support	0.068	0.050	0.013	0.023	0.000	0.041	0.026	0.023
Vocational-track <i>Hauptschule</i> ^a					-0.401***	0.098	0.043***	0.078
Academic-track <i>Gymnasium</i> ^a					0.517***	0.093	-0.105***	0.067
R^2_{between}		0.360		0.375		0.710		0.485

Note: ^a dummy-coded, reference category: intermediate-track *Realschule*.

B = unstandardized path coefficients, SE = standardized error.

***: $p < .001$; *: $p < .05$.