Learning and Instruction 17 (2007) 372–388

Learning and Instruction

www.elsevier.com/locate/learninstruc

The homework–achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort

Ulrich Trautwein*

Center for Educational Research, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

Abstract

The popular claim that homework time is positively related to achievement and achievement gains was tested in three studies. Time on homework was compared and contrasted with other indicators of homework assignment (i.e., homework frequency) and students' homework behavior (i.e., homework effort). The results of the three studies indicate that homework assignments are positively associated with achievement (class-level effect) and that doing homework is associated with achievement gains (student-level effect), but that the positive effects of homework assignments and completion are not captured by the ''time on homework'' measure.

Keywords: Homework; Time on task; Mathematics achievement; Multilevel modeling

In most countries around the world, student homework contributes substantially to time on task in core subjects. The 2000 cycle of the Programme for International Student Assessment (PISA), which was initiated by the Organisation for Economic Co-operation and Development (OECD; OECD, 2001) and obtained homework data from 32 countries, established that the reported weekly time spent on homework in mathematics, the language of assessment, and natural sciences ranged from 2.9 (Japan) to 7.0 (Greece) h, with countries such as the United States (4.6 h) and Germany (4.5 h) being close to the overall country mean. There is a widespread belief that time on homework is associated with greater achievement gains (Cooper, 1989; Cooper, Robinson, & Patall, 2006; Hattie & Clinton, 2001; Walberg & Paschal, 1995), but empirical support for the homework–achievement relation is not unequivocal (e.g., Cooper, Lindsay, Nye, & Greathouse, 1998; De Jong, Westerhof, & Creemers, 2000; see Trautwein & Köller, 2003a). The present article presents results from three studies permitting a detailed investigation of the relationship between homework and achievement.

1. The homework–achievement relationship

Homework is a ''complicated thing'' (Corno, 1996), a ''battlefield'' (Cooper, 2001) for teachers, students, and parents. Teachers complain about students failing to complete their assignments, students resent the time that homework

* Tel.: +49 (0)30 82406451; fax: +49 (0)30 82406490.
E-mail address: trautwein@mpib-berlin.mpg.de

0959-4752/$ - see front matter
doi:10.1016/j.learninstruc.2007.02.009

takes away from more enjoyable activities, and parents grumble about the family stress caused by disagreements on whether, when, and how to do homework (Burnett & Fanshawe, 1997; Cooper, 2001; Grolnick, 2003; Hoover-Dempsey et al., 2001; Warton, 2001). Yet despite these downsides, the majority of teachers, students, and parents remain convinced that homework is a valuable educational tool (e.g., Cooper et al., 1998; Xu, 2005). Reviews of homework research (Cooper, 1989; Cooper et al., 2006; Cooper & Valentine, 2001; Keith, 1986; Walberg, 1991) and, more generally, research on educational effectiveness (e.g., Hattie & Clinton, 2001) support the view that time spent on homework is associated with achievement gains. These reviews are frequently cited as documenting positive homework effects.

The majority of non-experimental studies on the homework–achievement relation (Cooper, 1989; Cooper et al., 2006; Paschal, Weinstein, & Walberg, 1984) have examined the impact that the time spent on homework has on achievement (also see Keith, Diamond-Hallam, & Goldenring Fine, 2004; OECD, 2001, for recent examples). However, the significance of many of these studies is debatable (see Cooper, 1989; Keith, 1986; Trautwein & Köller, 2003a). There are at least three potential threats to the validity of typical correlational studies on the homework–achievement relationship. First, homework can be related to achievement at two levels (see Trautwein & Köller, 2003a). One, a homework effect *at the class level* (or *homework assignment effect*) is found when students in classes with a higher quantity or quality of homework have more pronounced achievement gains than students in other classes (e.g., De Jong et al., 2000; Trautwein, Köller, Schmitz, & Baumert, 2002). The other, a homework effect *at the student level* (or *homework completion effect*), is found when students in the same class who differ in their homework behavior (e.g., time spent on homework) show differential outcomes (e.g., Cooper et al., 1998). In this sense, homework is a classic example of the multilevel problem (e.g., Kreft & de Leeuw, 1998; Raudenbush & Bryk, 2002), and it is essential to differentiate between teacher- and student-level effects in all studies that relate homework to achievement. Unfortunately, few studies have attended to this aspect (Trautwein & Köller, 2003a). Hence, class-level and student-level effects are confounded in many studies, and this precludes any firm conclusions about the effectiveness of homework assignments or homework completion.

A second major issue in several homework studies is that they do not control for the role of confounding variables. For instance, students who attend an advanced course in mathematics or an elite school might spend more time on mathematics homework than students enrolled in a basic course or a lower level school (see Keith & Cool, 1992). Likewise, more homework might be set in high-quality schools attended by students from privileged backgrounds. Similarly, female gender is known to be associated with higher achievement, more ambitious course-taking, and higher self-discipline during the high-school years (Duckworth & Seligman, 2006). Given these relationships, the finding of a positive relationship between homework and achievement might be attributable to a common cause (i.e., course level, school quality, or gender), and not to time on homework per se. Hence, the homework–achievement relationship might be spurious. In a related vein, the majority of studies are single-measurement studies. Questions pertaining to the directionality of homework effects cannot be readily answered on the basis of such designs. Does homework time affect later achievement or does achievement affect homework time – or do both effects exist? Longitudinal designs that control for prior knowledge and prior homework time are needed to address this question.

Third, research has concentrated almost exclusively on *time* spent on homework. This measure may in fact obscure the relationship between homework and achievement, rather than casting light on it. With reference to Carroll (1963), conscientious homework behavior is often equated with the time spent on homework. However, this perspective disregards the fact that Carroll's model predicts learning outcomes based on both time spent and time needed. Time needed is higher in students with low cognitive abilities and/or low prior achievement. Moreover, Carroll emphasized the role played by motivational and volitional factors (*perseverance*). In referring to time on task, Carroll in fact meant only the *active* time on task. Yet all sorts of distractions can have detrimental effects on students' homework behavior. If a student reports spending a lot of time on his or her homework, this is not necessarily a sign of great conscientiousness, but may reflect problems of motivation or concentration (see Trautwein & Köller, 2003a, for a critical account of the time on task variable).

Several recent studies that have separated the effects of homework assignment and homework completion (Cooper et al., 1998; De Jong et al., 2000; Muhlenbruck, Cooper, Nye, & Lindsay, 2000; Trautwein et al., 2002; Trautwein & Köller, 2003b; but see Cooper et al., 2006) indicate that students who spend more time on homework do not outperform their peers—in fact, some studies have shown these students to lag behind their peers in terms of achievement and achievement gains. At the class level, a higher number of homework tasks (De Jong et al., 2000) and higher homework frequency (Trautwein et al., 2002) have proved to be associated with higher achievement gains, but more time spent on homework has not.

It needs to be re-emphasized that time on task describes only one aspect of homework behavior. The *effort* a student invests in homework is also assumed to have a positive impact on achievement gains, but is not necessarily related to homework time (Schmitz & Skinner, 1993; Trautwein & Köller, 2003b; Trautwein, Lüdtke, Schnyder, & Niggli, 2006). For instance, Schnyder, Niggli, Cathomas, Trautwein, and Lüdtke (2006) used data from a sample of 1,832 students from 104 classes in three Swiss cantons to examine the effects of time and effort on homework in French as a foreign language. The students were assessed at the beginning and end of grade 8. To avoid confounding class-level and student-level homework effects, Schnyder et al. standardized ($M = 0$, $SD = 1$) all variables within school classes; hence, their study exclusively focused on homework completion (and not homework assignment) effects. Irrespective of the method used to tap time spent on French homework (global self-report, diary method) and of the dimensions examined (e.g., written work, learning vocabulary), homework time proved to be negatively related to achievement and achievement gains. By contrast, effort on homework was consistently associated with higher student achievement and greater achievement gains.

2. The present studies

In response to the problems identified with the popular "time on homework" measure and the typical design of homework studies (see Trautwein & Köller, 2003a), the present studies critically assess the relationship between homework variables and achievement. The view that time on homework is associated with achievement gains (Cooper, 1989) is challenged, and the neglect of the multilevel nature of homework that was shown in, for instance, the OECD (2001) report on the PISA 2000 study is criticized. However, I do not claim that homework has no positive effects on achievement. On the contrary, homework effects might in fact be larger and more consistent than has been argued previously.

Three data sets were analyzed. In the first study, the homework–achievement relationship reported by the OECD (2001) for the PISA 2000 data set was re-examined, using data from the large German extension to the study. Second, data from the Third International Mathematics and Science Study (TIMSS; Beaton et al., 1996) were used to contrast time on homework with homework frequency in a longitudinal design. In both studies, multilevel modeling was used to disentangle homework effects at the student level and the class/school level. Third, the reciprocal relationship between students' effort on homework and their achievement was analyzed in more detail using a longitudinal data set tailored to homework research. Because all data sets include mathematics homework variables and mathematics achievement indicators, the analyses presented here focus on mathematics.

3. Study 1

The large-scale Programme for International Student Assessment (PISA; OECD, 2001) provides internationally comparable data on the schools systems of countries belonging to the Organisation for Economic Co-Operation and Development (OECD) and several non-OECD countries. The PISA study has attracted widespread attention around the world and has influenced national education policies in many countries.

One finding of the first PISA cycle in 2000 was evidence for a positive relationship between homework and achievement in the vast majority of participating countries. In PISA 2000, a four-point scale with the response categories "no time," "less than 1 h a week," "between 1 and 3 h a week," and "3 h or more a week" was used to assess time on homework in three subjects (language of assessment, mathematics, science). Based on these data, an index of total time spent on homework was constructed. This combined scale was then related to achievement in reading. The PISA tests were scaled to have an international mean of 500 and a standard deviation of 100.

Analyses using this combined homework index revealed an impressive relationship between homework and achievement (see OECD, 2001, p. 300, Table 7.6). In the United States, for instance, there was an increase in reading achievement of 26.53 ($SE = 1.97$) per homework unit. Thus, for every unit increase on the combined homework scale, student achievement increased by about a quarter of a standard deviation. For Germany, the reported achievement gain was lower, but still notable (achievement gain = 11.57 per homework unit; $SE = 2.07$). On average across the OECD countries, an increase of $M = 13.87$ ($SE = .44$) in reading achievement was registered with every additional homework unit. The combined homework index explained 4.5% of the variation in student performance in reading across the OECD countries.

The PISA study implemented rigorous standards in the development of its tests and used modern data analysis techniques to compare the educational output of the participating countries. Hence, at first glance, the PISA data set seems especially well suited to studying homework effects in countries around the world. Nevertheless, the PISA results on homework should be approached with caution for at least three reasons. First, and most important, no distinction was made between homework assignment effects and homework completion effects. Hence, the precise nature of the relations between homework and achievement at the student level and the class level remains unclear. Second, a composite homework index comprising language, mathematics, and science was used to predict *reading* achievement. Yet there is little reason to believe that mathematics or science homework will have a strong impact on reading outcomes. Third, no control was made for prior knowledge, basic cognitive abilities, or differences in school quality. These variables may affect both time on homework and achievement (e.g., more homework might have been assigned in schools that cater to high-achieving students).

Study 1 is a re-analysis of the PISA 2000 findings using data from the German extension to the study, three major features of which permit a critical investigation of the homework–achievement relationship. First, in addition to the PISA achievement measures (reading, mathematics, science), a test of cognitive abilities was administered. Hence, it is possible to control for the effects of cognitive abilities when studying the homework–achievement relationship. Second, whereas the international OECD study tested 15-year-old students, the extended German sample included a large additional sample of ninth graders. Hence, any bias in the homework–achievement relationship caused by students being enrolled in different grade levels can be controlled. Third, a national mathematics test with high curricular validity (Klieme, Neubrand, & Lüdtke, 2001) was administered to all participating students, in addition to the international tests. Should homework effects exist, they are more likely to be statistically detected by a test with high curricular validity.

The research questions addressed in Study 1 were as follows. First, how much of the variance in time on homework is between classes and how much is between students? In other words, should homework time be treated as a class-level variable, a student-level variable, or a combination of both? Are school type, cognitive abilities, and gender significant predictors of homework time? Second, is homework time at the class level and the student level related to achievement once school type, cognitive abilities, and gender (variables that have been omitted from many studies) are controlled?

3.1. Method

3.1.1. Sample

This study is based on data from the German extension (Baumert et al., 2002) to the PISA 2000 study (OECD, 2001). The main goal of the extension was to permit comparisons between the 16 German states by drawing an expanded, nationally representative sample. A multistage sampling procedure was implemented to ensure high representativity of the data, and participation rates of 85% and above were achieved, with a total of 34,765 grade 9 students. The analyses reported here are restricted to the 24,273 students attending one of the three major school types in Germany: *Hauptschule* (lower track; 22.9%), *Realschule* (middle track; 34.4%), *Gymnasium* (highest track; 42.7%). Students in mixed track schools, comprehensive schools, and vocational schools were excluded because of the difficulty in accounting for course-level effects in these schools. 50.9% of the participating students were female, and the mean age was 15.70 (SD = .57). The extension study was conducted the day after the international PISA instruments had been administered.

3.1.2. Instruments

3.1.2.1. Mathematics achievement test. The mathematics test implemented in the German extension to PISA 2000 had high curricular validity, having been successfully tailored (see Klieme et al., 2001) to students in German schools, and was administered to all students in the sample. The correlation between students' scores on the German national mathematics assessment and the international mathematics assessment was $r = .85$. Mathematics scores for individual students were generated using item response theory techniques. The resulting test score distribution had a mean of $M = 100$ (SD = 30); the reliability of the test was .89 (formula by Rost, 1996).

3.1.2.2. Basic cognitive abilities. The *Figure Analogies* and *Verbal Analogies* subscales from the Cognitive Abilities Test 4-12+R (Heller & Perleth, 2000), a German version of the CogAT by Thorndike and Hagen (1993), were used to tap *cognitive abilities*. The scales consist of 25 figural items and 20 verbal analogy items in multiple-choice format and are considered to constitute a test of *reasoning* that is relatively free of environmental effects. A composite score with $M = 50$ (SD = 10) was generated using item response theory techniques.

3.1.2.3. Homework time. The original PISA item assessing time on homework *in mathematics* was implemented, using a four-point scale with response categories "no time" (coded 1), "less than 1 h a week" (2), "between 1 and 3 h a week" (3), "3 h or more a week" (4).

3.1.2.4. School type. The German school system is characterized by its early and selective assignment of students to different *school types*. Tracking takes place after grade 4 (after grade 6 in a few states), when students are aged about 10. Students are typically assigned to one of three different types of secondary school: *Hauptschule*, the academically least demanding track; *Realschule*, the intermediate track, and *Gymnasium*, the highest track. Dummy variables were created to examine the effects of school type. The middle track (*Realschule*) was used as the reference category.

3.1.2.5. German states (Bundesländer). All 16 German states participated in the German extension to PISA 2000, and systematic differences were found between the states (Baumert, Trautwein, & Artelt, 2003). A dummy variable coding scheme was applied to control for these effects by taking into account differences between the states in later regression analyses.

3.1.3. Statistical analyses

Multilevel regression analyses were conducted to predict homework time and achievement. In most studies conducted in school settings, individual student characteristics are confounded with classroom or school characteristics because individuals are not randomly assigned to groups. This clustering effect introduces problems related to appropriate levels of analysis, aggregation bias, and heterogeneity of regression. Moreover, problems of model mis-specification arise due to the lack of independence between measurements at different levels.

For the present research, it is particularly important to note that the meaning of a variable at the student level may not bear any straightforward relation to its meaning at the classroom level. This holds particularly for the "time on homework" measure, which provides different information at the individual versus class level. The average time typically spent on homework at the class level is a proxy measure of the typical assignment length in the class, whereas the time typically spent on a homework assignment at the student level might either signify a student's working speed or the effort he or she makes to complete homework assignments.

Particularly when major variables represent different levels, it is important to analyze data with appropriate multilevel statistical procedures. Multilevel modeling, a special form of regression analysis, provides a powerful methodology for handling hierarchical data, and was used in this study. A detailed presentation of multilevel modeling (also referred to as hierarchical linear modeling, HLM) is beyond the scope of the present investigation and is available elsewhere (e.g., Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). In the present study, students were used as level 1 (or student-level) variables, and schools as level 2 (or school-level) units. The HLM 5 software (Raudenbush, Bryk, Cheong, & Congdon, 2000) was used for the analyses. Because HLM does not report standardized regression coefficients, the original metric of all variables was conserved in the analyses reported. All student-level variables were entered uncentered.

Missing data are inevitably a major challenge in large-scale studies, particularly if the proportion of missing data exceeds 5% (Graham, Cumsille, & Elek-Fisk, 2003). For each of the items and scales considered here, the proportion of missing data was below 10%. To deal with missing data, the multiple imputation procedure (Schafer, 1997) was implemented. Using the NORM software (version 2.03, see Schafer & Graham, 2002) and several auxiliary variables (see Collins, Schafer, & Kam, 2001), five data sets were generated in which missing data were replaced with estimated values. The HLM 5 software (Raudenbush et al., 2000) was then used to simultaneously analyze all five imputed data sets. Descriptive reports are based on the first imputed data set.

3.2. Results

For time spent on mathematics homework, the most frequently endorsed category (43.7%) was ‘‘between 1 and 3 h a week,’’ followed by ‘‘less than 1 h a week’’ (33.7%). Another 16.1% reported spending ‘‘3 h or more a week’’ on their mathematics homework, and only 6.5% reported typically spending no time on mathematics homework at all. With a mean score of $M = 102.39$ (SD = 24.61) on the national mathematics test, the mean for this subsample corresponded closely to the overall national average of $M = 100$.

Two sets of multilevel regression analyses were performed. First, time spent on mathematics homework was used as the dependent variable, and several variables at the student (e.g., gender) and school (e.g., school type) levels were introduced as predictor variables. Second, these variables (including time spent on homework) were used to predict mathematics achievement. Table 1 reports the results of these analyses.

With time spent on homework as the dependent variable, a so-called empty model was calculated to estimate the amount of variance within and between schools (the so-called ‘‘variance components’’). It emerged that only 6.5% of the variance was between schools, and 93.5% was within schools. Hence, there were considerable differences in time spent on homework among students attending the same school. In a next step, a model was specified in which the impact of the different German states was controlled (Model 1). In other words, the 15 dummy variables representing the 16 German states were introduced as level 2 variables. This analysis pointed to some significant cross-state differences. About 21% of the variance between schools was explained by the inclusion of these dummy variables (see Model 1 in Table 1; because they are not relevant in the present context, the regression weights of the dummy variables representing the German states are not shown). As reported at the bottom of Table 1 (‘‘variance components’’), after controlling for the impact of the German states, an even more pronounced difference emerged between the variance located between students (level 1 variance = .62) and the variance located between schools (level 2 variance = .03).

Model 2 additionally included two student-level predictors (gender, cognitive abilities) and the school type dummy variables as level 2 predictors. At the student level, female students reported spending more time on homework than did male students. Furthermore, cognitive abilities were negatively related to time spent on homework. At the school level, relative to the reference category of *Realschule* (middle track), there was a positive effect of attending a *Gymnasium* (highest track) and a negative effect of attending a *Hauptschule* (lower track). The percentage of variance

Table 1
Predicting time spent on homework and mathematics achievement: results of multilevel modeling (Study 1)

	Time spent on homework			Mathematics achievement				
	Model 1[a]	Model 2[a]		Model 1[a]	Model 2[a]		Model 3[a]	
		B	SE		*B*	SE	*B*	SE
Level 2: schools								
School type (reference category: *Realschule*)								
Hauptschule (lowest track)		−.25***	.02				−9.16***	.37
Gymnasium (highest track)		.14***	.02				13.99***	.35
Homework time					24.67***	2.03	1.76**	.58
Level 1: students								
Cognitive abilities		−.01***	.00				1.30***	.01
Gender: female		.23***	.01				−1.50***	.11
Homework time					−3.45***	.15	−7.84***	.18
Variance explained								
Level 2	.21	.35		.19	.26		.97	
Level 1	.00	.03		.00	.03		.43	
Variance components								
Level 2	.03	.03		326.16	297.53		12.86	
Level 1	.62	.60		226.32	218.89		129.23	

Note. *B*: unstandardized regression coefficient. SE: standard error of *B*.
$^{***}p < .001$; $^{**}p < .01$.
[a] Fifteen dummy variables controlling for the effects of the 16 German states are included in all models reported in this table.

explained at level 2 by state and school type variables was substantial (35%); the percentage explained by gender and cognitive abilities was rather small (3%).

Taken together, the analyses indicate that a large amount of variance in the time spent on homework is located at the student level, highlighting the need to consider this aspect when examining the homework–achievement relationship. Moreover, there were significant effects of group membership at level 2 (states, school types). Neglecting these group differences when relating homework time to achievement might result in spurious effects. This is further demonstrated in the following analyses.

The second set of analyses analyzed the effects of homework time at the student and the school level on achievement. The variance in mathematics achievement within schools relative to between schools was estimated in a first, empty model (not reported in Table 1). It emerged that 64% of the variance was between schools, reflecting the marked differences in student achievement across the three school types covered in the present study. The model specified in the next step controlled for the differences in achievement levels across the 16 German states (Model 1). These differences accounted for 19% of the between-school variance.

In Model 2, homework time was included as a predictor at the student level (time spent on homework reported by individual students) and the school level (average time spent on homework reported by students in a given school). As documented in Table 1, homework time had a negative effect at the student level and a positive effect at the school level. In other words, students who spent more time on mathematics homework than their schoolmates scored lower on the mathematics assessment, whereas a high average homework time at the school level was positively related to achievement. Homework time explained 7% of the variance at level 2 and 3% at level 1. However, because potentially important predictor variables, such as cognitive abilities or gender, were not controlled in Model 2, the homework-achievement relationship identified might be spurious.

In Model 3, therefore, gender and cognitive abilities were included as additional level 1 variables and school type as an additional level 2 variable. All of these variables were statistically significantly related to mathematics achievement. Controlling for the other variables in the model, *Gymnasium* students achieved the best mathematics scores, followed by *Realschule* students. The difference between the highest and lowest track was equivalent to almost one standard deviation. Similarly, at the student level, cognitive ability was closely associated with mathematics achievement. The effect of gender was small (controlling for the other variables, boys scored 1.50 points higher than girls), but given the large sample size, the effect was statistically significant.

The effects of homework time are, of course, of particular interest in the present study. At the school level, when the additional predictor variables were included in Model 3, the size of the homework time effect decreased dramatically (from 24.67 in Model 2 to 1.76 in Model 3). Hence, homework time as a school-level variable was obviously closely related to other achievement-related variables (i.e., school type). When controlling for these variables, it only had moderate incremental predictive power.

Interestingly, the picture to emerge at the student level was a different one, with the predictive power of time spent on homework increasing when the other predictor variables were included (from −3.45 in Model 2 to −7.84 in Model 3). The negative sign of the beta coefficient indicates that—when controlling for the other variables in Model 3—the students who spent more time on their mathematics homework than their schoolmates scored lower on the mathematics test.

3.3. Discussion of Study 1

The PISA 2000 (OECD, 2001) findings that longer homework time is associated with higher achievement provided the starting point for Study 1, the results of which challenge this notion. Using the German extension of the PISA data set, potential common causes of homework time and achievement were controlled, and multilevel modeling was applied to disentangle student-level and school-level effects. The analyses indicate that the relationship between homework time and achievement is only moderate at the school level and is in fact negative at the individual level.

The large sample sizes and high-quality assessments of the PISA 2000 data set provide a unique opportunity to study homework effects, but the PISA design also has its limitations in the present context. First, the cross-sectional design implemented in PISA 2000 makes it impossible to determine the causal direction of the effects observed. Does it take low-achieving students more time to complete their homework because their prior knowledge is low, or might long homework times indicate that concentration difficulties result in low progress? Second, PISA 2000 used schools as sampling units, rather than school classes. Yet it is quite possible that the homework policies of different teachers in the same schools might vary. Hence, stronger homework effects might be found if school classes are used as sampling units. Finally, the homework measure used in PISA collapses time spent on homework and homework frequency. This might not be the optimal approach, given that homework frequency and homework time might be differentially associated with achievement (see Trautwein & Köller, 2003a).

4. Study 2

Study 2 was designed to address some of the limitations noted in Study 1. The empirical basis for this investigation was provided by a German sample of lower secondary students who participated in the Third International Mathematics and Science Study TIMSS (Baumert et al., 1997; Beaton et al., 1996). Again, there was a German national extension to the international design of TIMSS—most importantly, a longitudinal component was added. One year before they officially participated in TIMSS in grade 8, the mathematics achievement of most of these students was assessed in grade 7. Furthermore, one intact class was sampled in each of the participating schools (students did not switch classes between measurements). Finally, several additional instruments were administered in the national extension; e.g., homework time and frequency of homework assignments. Hence, by using a longitudinal data set with classes (instead of schools) as sampling units and separating homework frequency from homework time, Study 2 promises clearer insights into the potential effects of homework.

Again, two broad sets of research questions were studied. In the first set of analyses, homework time was used as the dependent variable. What percentage of the total variance is located at the class level? Are school type, cognitive abilities, and gender again significant predictors of homework time? Is prior achievement significantly related to time on homework? In the second set of analyses, mathematics achievement in grade 8 was the dependent variable, and the impact of homework variables (homework frequency and homework time at class and student level) was tested with and without controlling for potentially confounding variables (school type, cognitive abilities, prior achievement, gender). The major question to be addressed here was whether homework still predicts achievement gains when other important predictor variables are included in the model.

4.1. Method

4.1.1. Sample

The German longitudinal extension to TIMSS was based upon a total of 2939 students from 120 classes. For the present study, 20 classes in which the homework instruments were not administered were dropped. Moreover, as in Study 1, students attending comprehensive schools (nine classes) were not included. The resulting final sample consisted of 2216 students (49.5% female, mean age in grade 7 = 13.7 years) from 91 classrooms, and included students from three different school types: *Hauptschule* (26.7%), *Realschule* (34.9%), and *Gymnasium* (38.4%). Nearly all (95%) participants were German. Active consent was collected from all parents. Trained research assistants administered the instruments during regular school hours.

4.1.2. Instruments

4.1.2.1. Mathematics achievement. The 36 mathematics items administered in grade 7 (T1) were taken from previous studies run by the International Association for the Evaluation of Educational Achievement (IEA), in particular, the First and Second International Mathematics Study (Husén, 1967; Robitaille & Garden, 1989), and from an earlier investigation conducted by the Max Planck Institute for Human Development (see Baumert, Roeder, Sang, & Schmitz, 1986). Achievement at grade 8 (T2) was measured by a total of 158 items from the official TIMSS item pool. The items were distributed over eight booklets. Each booklet contained between 30 and 40 items, some specific to that booklet, and some anchor items common to all booklets. Students worked on one booklet each, thus allowing a broad range of subject matter to be covered without exhausting the students. All items were checked for curricular validity within each country. Responses were scaled using item response theory techniques (for details of the item contents and scaling procedures, see Martin & Kelly, 1997). Because a selection of T1 items was repeated at T2, it was possible to construct a common achievement metric for T1 and T2 using item response theory (IRT; for details, see Baumert et al., 1997). Although the content areas and performance categories of the items varied, analyses based on Item Response Theory (IRT) revealed that a unidimensional model was appropriate for describing the latent variable underlying the test results at both measurement points. The coefficient alpha estimates of reliability based on achievement data used in the present investigation were .81 and .88 for responses in grades 7 and 8, respectively. To enhance the interpretability of the results of the present study, achievement at T1 was standardized ($M = 0$, SD = 1). The achievement score at T2 was rescaled according to the metric of the first measurement point.

4.1.2.2. Basic cognitive abilities. The *Figure Analogies* and *Verbal Analogies* subscales from the Cognitive Abilities Test were again used to tap *cognitive abilities*. Internal consistency was high (Kuder-Richardson formula 20 = .92), and a sum score was generated ($M = 28.62$, SD = 9.81).

4.1.2.3. Homework variables. One question tapped the *frequency of homework* assignments: "How often are you usually assigned mathematics homework?" Response categories ranged from *never* (coded 1) to *always* (coded 5). Another question investigated the *time a student typically spent on an assignment* when homework was given: "How long does it usually take you to finish your mathematics homework?" Response categories were *15 min* (coded 1), *30 min* (2), *60 min* (3), *90 min* (4), *120 min* (5), and *more than 120 min* (6). Time typically needed for homework was considered at both the individual level and the class level (see SECTION 4.1.3). Homework variables were collected at T2.

4.1.3. Statistical analyses

Again, hierarchical linear modeling was applied. Homework time was first split into a between-class and a within-class components. In Germany, homework is typically given to all students in a class. Therefore, the frequency of homework assignments within a class is constant, and does not need to be considered at the individual level (see Raudenbush & Bryk, 2002). As an indicator of the reliability of the class-average frequency variables, the intraclass correlation coefficients for homework frequency were calculated (see Bliese, 2000; Snijders & Bosker, 1999). The ICC(1)—which indicates the homogeneity of the frequency ratings within classes—was .36. With an average of 24 students per class, the reliability of the class-averages frequency score—the so-called ICC(2) —was .93; thus, the class-average homework frequency variable was very reliable. As in Study 1, the original metric of the variables was retained, and level 1 predictor variables were introduced uncentered.

As mentioned above, missing values are an unavoidable challenge in large-scale and longitudinal studies. In the present study, various students had missing data on single items (item nonresponse) or at one of the measurement points for a variety of reasons including sickness, lack of written parental consent, technical errors in the questionnaire, and refusal to answer certain items. For the variables used in the present study, an average of 13.6% of the data was missing, with a maximum of 24.5% missing values for cognitive abilities. In line with current thinking on missing data (e.g., Graham et al., 2003), we again implemented the multiple imputation approach. Five imputed data sets were generated using the NORM software (version 2.03, see Schafer & Graham, 2002). The HLM 5 software (Raudenbush et al., 2000) was then used to simultaneously analyze all five imputed data sets.

4.2. Results

On average, homework was assigned frequently ($M = 3.92$, SD = 1.07) in the present sample. With regards to homework time, the most frequently endorsed category (44.9%) was "30 min," followed by "15 min" (23.6%), and "60 min" (21.4%). At the class level, there was a correlation of .27 ($p < .01$) between homework time and homework frequency (student level: $r = .24$, $p < .001$). This indicates that homework time and homework frequency are two clearly separable aspects of homework. Mathematics achievement increased by more than half a standard deviation between T1 ($M = .00$, SD = 1.00) and T2 ($M = .62$, SD = 1.14).

Again, two sets of multilevel regression analyses were performed. Time spent on mathematics homework was used as the first dependent variable (see Models 1 and 2 in Table 2). The empty Model 1 indicated that only 4% of the variance in time spent on homework was located between classes, meaning that there was much more variation in homework time among different students in the same class than among different classes. In Model 2, two predictor variables were significantly related to homework time. On average, controlling for the other predictor variables, homework time was higher in *Gymnasium* classes than in the reference group (*Realschule* classes). Moreover, students with high T1 mathematics achievement reported spending less time on homework than their lower achieving classmates. Cognitive abilities and gender were not significantly related to time on homework. Hence, this first set of analyses indicates

Table 2
Predicting time spent on homework and mathematics achievement: results of multilevel modeling (Study 2)

	Time spent on homework			T2 mathematics achievement				
	Model 1	Model 2		Model 1	Model 2		Model 3	
		B	SE		*B*	SE	*B*	SE
Level 2: classes								
School type (reference category: *Realschule*)								
Hauptschule (lowest track)		−.04	.10				−.30**	.08
Gymnasium (highest track)		.33**	.10				.54***	.09
Homework frequency					.59***	.12	.11*	.05
Homework time					.10	.28	−.02	.12
Level 1: students								
Cognitive abilities		.00	.00				.02***	.00
T1 math achievement		−.15***	.04				.42***	.03
Gender: female		−.10	.06				−.09**	.03
Homework time					−.13***	.02	−.08***	.02
Variance explained								
Level 2		.00			.20		.88	
Level 1		.02			.03		.31	
Variance components								
Level 2	.06	.06		.72	.58		.08	
Level 1	1.23	1.21		.58	.56		.40	

Note. *B*: unstandardized regression coefficient. SE: standard error of *B*.
$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$.

a simultaneous, counterbalancing effect: whereas more homework was assigned to high-achieving classes (*Gymnasium* effect at level 2), the higher achieving students within each class took less time to complete their assignments.

In the second set of multilevel analyses, T2 mathematics achievement was used as the dependent variable. The empty Model 1 identified considerable differences in achievement between the participating school classes, accounting for 55% of the total variance. These substantial differences are a consequence of the high level of selectivity in the German school system. In Model 2, the three homework variables were introduced. At the class level, more frequent homework assignments were associated with higher achievement at T2, but the coefficient for homework time was not significant. Homework time at the student level negatively predicted achievement.

In the next step (Model 3), cognitive abilities, gender, and school type were introduced. Moreover, T1 mathematics achievement was included as an additional predictor variable, allowing the effects of all other predictor variables to be interpreted as effects on mathematics achievement at T2, controlling for prior math achievement. As expected, T1 mathematics achievement was a powerful predictor of T2 mathematics achievement. At the class level, the largest achievement gains were found in the highest track (*Gymnasium*), followed by the middle track. The direct effect of homework frequency was much smaller than in Model 1 (dropping from .59 to .11), but it remained significant. Homework time was not significantly related to achievement gains at the class level. At the student level, achievement gains were more pronounced in male students, students with high cognitive abilities, and students who spent *less* time on homework than other students. The amount of variance explained by the variables included in Model 2 was high (88%) at the class level and substantial (31%) at the student level.

4.3. Discussion of Study 2

Using a longitudinal data set with classes (instead of schools) as sampling units and distinguishing homework frequency from homework time, Study 2 yielded several important results. First, homework frequency—but not homework time—was a significant predictor of achievement at the class level. The homework measure used in Study 1 (time on homework per *week*) combined homework time and homework frequency, but the present study indicates that these two aspects of homework should be studied separately. Second, homework time was negatively related to achievement and achievement gains at the student level. This indicates that lengthy homework times are unlikely to reflect student motivation or effort, but rather indicate motivational problems or problems of understanding. Third, as documented by the change in the regression coefficients of the homework variables across the different models, when failing to control for important additional predictor variables such as school type, cognitive abilities, or prior knowledge, the effects of homework variables are artificially inflated.

It is widely accepted among teachers, students, and parents that homework is associated with enhanced achievement (e.g., Cooper, 1989; Grolnick, 2003; Haag, 1991). Somewhat surprisingly, then, in both Study 1 and Study 2, homework time at the individual student level was negatively related to achievement. However, time on homework does not necessarily equate with effort on homework (see Trautwein & Köller, 2003a). Unfortunately, due to its restricted set of variables, the TIMSS data set does not allow the exact relationship between *time on homework* and *effort on homework* to be studied. Hence, Study 3 was designed to permit an in-depth examination of this relationship.

5. Study 3

Study 3 examined the effects of students' homework behavior on their achievement. To this end, homework behavior was split into two components: homework effort (i.e., conscientious execution of homework assignments) and homework time. The longitudinal data set used was especially tailored to address questions pertaining to the utility of homework (see Trautwein & Kropf, 2004). Path analysis was used to test the reciprocal relationships between homework and achievement. More specifically, the three dependent variables—T2 achievement, T2 homework effort, and T2 homework time—were simultaneously predicted by the corresponding T1 variables as well as by gender and cognitive abilities.

Three main research questions were addressed. First, how close is the relationship between effort invested in homework and time spent on homework? Second, does effort on homework positively predict achievement gains when controlling for the other predictor variables? Third, is there again a nonsignificant or negative relationship between time on homework and achievement gains?

5.1. Method

5.1.1. Sample

A total of 483 eighth graders (58.8% female; $M = 13.45$ years of age, $SD = .58$) from 20 classes participated in this study. All students were attending state *Gymnasium* schools in Berlin, Germany. Student participation was voluntary, and written consent was collected from all parents. The study was conducted during regular school hours in intact classes in November (T1) and May (T2) of the 2003/2004 school year. Trained research assistants collected data in the participating classes.

5.1.2. Instruments

5.1.2.1. Achievement. Two sorts of achievement measures were used. First, at both measurement points, students reported their grade on their most recent mathematics test. These tests typically assess student knowledge of the material covered in the previous 8–10 weeks and are devised, administered, and corrected by the mathematics teacher. Second, students reported the mathematics grades they had been awarded on their report cards at the end of grade 7 and in the middle of grade 8. In Germany, these grades reflect the student's overall achievement in one semester. They are based on test results as well as on effort in class and at home. In German secondary schools, grades and test scores are assigned on a metric ranging from 1 (very good) to 6 (unsatisfactory). For the subsequent analyses, both achievement indicators were recoded so that high scores reflect high achievement.

5.1.2.2. Cognitive abilities. The *Figure Analogies* subscale from the Cognitive Ability Test 4-12+R (Heller & Perleth, 2000) was used to tap *cognitive abilities*. The internal consistency (Kuder-Richardson formula 20) of the cognitive ability test was .90.

5.1.2.3. Homework effort. Homework effort was measured by six items tapping homework compliance and effort invested in homework. Example items are "I do my best in my mathematics homework" and "I often copy mathematics homework from others" (reverse scored). Students high on homework effort complete their homework assignments carefully and do not copy from others. Internal consistency (Cronbach's alpha) was good for T1 (.86) and T2 (.84).

5.1.2.4. Homework time. At both time points, homework time was measured by the following question: "On average, how many minutes do you spend on each mathematics homework assignment you are set?" Students were asked to enter the number of minutes in a blank field.

5.1.3. Statistical analyses

Study 3 focuses on the effects of homework *completion*, rather than homework *assignment*. Accordingly, variation in change in achievement and homework behavior *within* (and not between) the classrooms is of central interest. Trautwein and Köller (2003a) proposed that achievement and homework variables be centered or standardized if the primary focus is on student-level effects (also see Corno, 1980). Following this recommendation, in the present study, all variables of interest were standardized ($M = 0$, SD = 1) within classes.

Simultaneous regression analyses were specified using Mplus 3.1 (Muthén & Muthén, 1998–2004), in which gender, cognitive abilities, T1 homework effort, T1 homework time, and one of the T1 achievement indicators were used to predict T2 achievement, T2 homework effort, and T2 homework time. Given the high intercorrelations between test scores and school grades at both measurement points, separate models were estimated for these two variables to avoid any unwanted effects of multicollinearity. As in Studies 1 and 2, only direct effects were estimated. Because all T1 correlations and T2 residual correlations as well as all possible regression coefficients were freely estimated, the models estimated were saturated. Hence, model fit cannot be computed.[1] The missing values estimator implemented in Mplus was used to deal with missing values (which ranged from 10% to 23% across the variables used). Mplus applies a model-based approach to missing data, which builds on a full information maximum likelihood estimation (see Allison, 2001).

5.2. Results

Zero-order correlations among the variables included in this study are shown in Table 3. Gender was significantly related to homework time, with girls reporting spending more time on mathematics homework than boys. Cognitive abilities were most closely related to mathematics achievement. T1 and T2 mathematics school grades and T1 and T2 scores in the most recent mathematics test were substantially correlated, with correlations as high as $r = .68$ ($p < .001$). Homework effort was significantly positively related to achievement, whereas the correlation with homework time was nonsignificant or negative. Homework time was significantly negatively related to achievement, both within and across the two measurement points.

Thus, as expected, homework time was associated with lower achievement, whereas homework effort was related to higher achievement. However, the zero-order correlations presented thus far cannot rule out the possibility that more time spent on homework is associated with higher achievement gains over time. Hence, in a second step, two models were specified in which T2 variables were regressed on T1 variables (see Table 4). Due to their high intercorrelation, separate models were specified for the two achievement indicators. Mathematics school grades (but not test scores) were entered in Model 1; test scores (but not school grades) were entered in Model 2.

In Model 1, T2 mathematics school grades, T2 homework effort, and T2 homework time were regressed on all three corresponding T1 variables as well as on gender and cognitive abilities. Table 4 shows the results (fully standardized regression weights) of this analysis. T1 school grades had the strongest predictive power with respect to T2 mathematics school grades. Controlling for T1 school grades, homework effort at T1 positively and statistically significantly predicted school grades at T2, whereas a high amount of time spent on homework at T1 was associated with *lower* achievement at T2. Only T1 homework effort and gender had a significant predictive effect on homework effort at T2, with girls increasing their homework effort relative to boys. Finally, T2 homework time was significantly related

[1] I also specified a second set of models in which homework effort was modeled as a latent variable. The fit of these models was acceptable (achievement test: χ^2 [105] = 326.397, RMSEA = .066, SRMR = .054; school grades: χ^2 [105] = 344.95, RMSEA = .069, SRMR = .056). The resulting regression coefficients for the homework effort variables increased only marginally when using the latent variable framework, and no effect on any other relationships was found.

Table 3
Zero-order correlations between gender, cognitive abilities, achievement, and homework behavior (Study 3)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) Gender: female	1									
(2) Cognitive abilities	.10	1								
(3) T1 mathematics school grade	.02	.25***	1							
(4) T1 mathematics test	.03	.32***	.65***	1						
(5) T1 homework effort	−.03	.11*	.42***	.40***	1					
(6) T1 homework time	.14**	−.13*	−.24***	−.30***	−.12*	1				
(7) T2 mathematics school grade	.01	.19***	.65***	.68***	.35***	−.19***	1			
(8) T2 mathematics test	.02	.23***	.54***	.53***	.40***	−.29***	.68***	1		
(9) T2 homework effort	.07	.03	.25***	.27***	.56***	−.12*	.37***	.32***	1	
(10) T2 homework time	.15**	−.05	−.20***	−.23***	−.04	.36***	−.24***	−.27***	−.06	1

Note. All variables except gender were standardized within classes prior to data analysis.
*** $p < .001$; ** $p < .01$; * $p < .05$.

to T1 homework time, gender, and T1 school grades. The effect of female gender indicates that, relative to boys, girls increased their time on homework from T1 to T2. Good school grades at T1 had a negative effect on T2 homework time, indicating that – when controlling for T1 homework time – high-achieving students needed less time for their homework at T2.

In Model 2, school grades were substituted by test scores. However, the pattern of results was almost identical to that reported for school grades. Most importantly, controlling for prior achievement, T2 achievement was positively predicted by homework effort at T1, and, controlling for T1 homework time, high T1 achievement predicted less time on homework at T2. The only major difference between Models 1 and 2 was the nonsignificant effect of T1 homework time on achievement at T2 in Model 2.

5.3. Discussion of Study 3

Study 3 indicates that homework effort is indeed positively related to achievement and achievement gains. The positive effect of doing homework was observed for both mathematics grades and mathematics test scores, and the beta weights were of substantial size. At the same time, homework time was not related—or negatively related —to achievement and achievement gains, thus corroborating the findings of Studies 1 and 2. Moreover, there was no positive correlation between homework effort and homework time. Hence, the present findings indicate that homework time is not a suitable indicator of the effort that students put into their homework.

Table 4
Predicting T2 school grades, T2 achievement, T2 homework effort, and T2 homework time: results (standardized regression coefficients) of longitudinal path analyses (Study 3)

Predictor variables	Model 1: school grade			Model 2: achievement tests		
	T2 school grade	T2 homework effort	T2 homework time	T2 achievement test	T2 homework effort	T2 homework time
Female	.01	.10*	.11*	.00	.09*	.12*
Cognitive abilities	.03	−.05	.02	.06	−.06	.01
T1 school grade	.54***	.02	−.15**	(−)	(−)	(−)
T1 achievement test	(−)	(−)	(−)	.43***	.06	−.17**
T1 homework effort	.16***	.56***	.06	.17***	.54***	.07
T1 homework time	−.14***	−.07	.32***	−.03	−.06	.31***

Note. (−) = Variable was not included as a predictor in this Model.
*** $p < .001$; ** $p < .01$; * $p < .05$.

6. General discussion

Three studies tested the popular claim that homework time is positively related to achievement and achievement gains. Time on homework was compared and contrasted with other indicators of homework assignment (i.e., homework frequency) and students' homework behavior (i.e., homework effort). The results indicate that homework assignments are positively associated with achievement (class-level effect) and that doing homework is indeed associated with achievement gains (student-level effect), but that the positive effects of homework assignments and completion are not captured by the "time on homework" measure.

6.1. The homework–achievement relation reconsidered

The central message of the present study is two-fold. First, researchers who study the homework–achievement relation are urged to consider the multilevel structure of homework and to control for important confounding variables. Second, there is indeed a positive, meaningful association between homework and achievement at both the class level and the student level, but more homework variables than time on homework need to be inspected to fully understand this association.

At the class level, homework assignments had a positive impact on achievement in both Study 1 and Study 2. Mirroring the impressive homework effects reported by the PISA study (OECD, 2001), the homework effects observed here were remarkably strong. In Study 1, for instance, an increase of one homework unit was associated with an increase in achievement of roughly one standard deviation. However, when other important predictor variables such as school type and cognitive abilities were controlled, the class-level homework effect was dramatically reduced. Variables such as these might affect both time on homework and achievement. Unfortunately, neither the OECD (2001) analyses nor several studies reviewed by Cooper (1989) controlled for potentially confounding variables. Hence, these studies' findings on the homework–achievement relationship should be approached with care.

In Study 1, homework time and homework frequency are confounded in the "time on homework" measure. When these two variables were separated in Study 2, there was evidence for a positive effect of frequency—but not of time—on change in achievement. Although the present study does not allow this effect to be explained in detail, it corresponds with the emphasis that several theories of learning and instruction place on regular, step-by-step learning opportunities (e.g., Weinert & Helmke, 1995). Likewise, Trautwein and Köller (2003a) have speculated that lengthy homework assignments may lead students to internalize wrong routines.

At the student level, the present study supports the view that effort on homework is associated with positive development in achievement. At the same time, as documented in Studies 2 and 3, the length of homework times is predicted by low prior achievement and does not positively impact later achievement, even when controlling for several other variables. This association between longer homework time and low achievement is quite plausible: it takes weaker students longer to complete a given set of tasks. However, the nonsignificant or even negative effect of homework time on achievement after controlling for other variables warrants some explanation. A first possible explanation is that spending a lot of time on homework might signify a rather inefficient, unmotivated homework style. Congruent with this view, homework time was only loosely (and negatively) related to reported homework effort. A second potential explanation is that the reliability of the "time on homework" measure is questionable (see Trautwein & Köller, 2003a). Indeed, De Jong et al. (2000) reported a rather low correlation between an overall "time on homework" measure and time on homework as documented in a logbook. Moreover, in the present study there was only a moderate correlation between T1 and T2 time on homework.

A third explanation focuses on the difference between homework time and other out-of-school learning activities (e.g., Cooper, Valentine, Nye, & Lindsay, 1999). Importantly, the homework measures used in the present study clearly targeted homework assigned by the teacher. Extra study time was not included. It is quite likely that "homework" measures covering various other out-of-school learning opportunities will find a more positive relationship between learning time and achievement.

6.2. Limitations and future research

With its large samples, longitudinal data, and converging effects across the three studies, the present article makes a strong case for the re-conceptualization of homework effects and homework research. At the same time, the present research has certain limitations. For instance, all three studies sampled students from a restricted age range (eighth and ninth graders). It is quite possible (see Muhlenbruck et al., 2000) that the homework–achievement relationship is stronger in upper grades than in lower grades.

Moreover, all samples were drawn from German schools. The first two studies used representative samples of German students; the third used a sample of students from 20 *Gymnasium* (highest track) classes in one city. Since all three studies contradict the popular claim that homework time is beneficial to student achievement (Cooper, 1989; Cooper et al., 2006), the question arises as to whether the results reported are country specific (Studies 1 to 3) and/or sample specific (Study 3). Of course, a comprehensive answer to this question would require a cross-cultural longitudinal study focusing on homework assignment and homework completion effects. Because no such study exists (Cooper et al., 2006), the evidence available needs to be carefully re-examined. Two issues seem to be of specific importance in this context.

First, the conceptual framework of the present research is worth highlighting. The studies took into account the multilevel nature of homework, controlled for potentially important additional predictor variables, and used longitudinal data (Studies 2 and 3). It is important to re-emphasize that a positive homework–achievement relationship was found for Germany in the PISA report (OECD, 2001). It was only after taking the hierarchical structure of the data into account and controlling for possible common causes such as school type that the effect at the individual level proved to be negative. Of course, some of the effects of additional predictor variables found for German schools might not apply in other school systems, or they may take different forms. For instance, although few school systems are as strictly differentiated as the German one, it is quite likely that effects of ability grouping will also emerge in countries such as the United States (e.g., Keith & Cool, 1992).

Second, focusing on the empirical evidence, several studies from other countries and with other German samples have yielded similar results. When controlling for the hierarchical structure of the data sets, recent studies from the Netherlands (De Jong et al., 2000) and Switzerland (Schnyder et al., 2006) have also reported negative effects or no effects of homework time on achievement gains. Although Cooper et al. (2006) reported a positive bivariate correlation between time on homework and achievement for 50 of the 69 correlations reported in studies of U.S. students conducted between 1987 and 2003, most of these studies did not take into account the hierarchical structure of the data. Hence, they are not comparable with the data reported in the present studies. When focusing on the relationship between homework time and achievement at the student level, Muhlenbruck et al. (2000) found no support for a positive relation between homework time and achievement; the average within-class correlations between student reports of time spent on homework and achievement amounted to $r = -.13$, ns, for elementary schools and $r = -.03$, ns, for secondary level. In Germany, other studies have reported a negative effect of homework time and a positive effect of homework effort on achievement (Schmitz & Skinner, 1999; Trautwein & Köller, 2003b).

Taken together, it is unlikely that the results reported in the present research can be fully explained by country-specific or sample-specific characteristics. Nevertheless, more research on this issue is clearly warranted.

It should also be noted that the present study was restricted to mathematics homework. There are some, albeit weak, indications (see Cooper, 1989; Cooper et al., 2006) that homework time is differentially related to achievement across different school subjects.

The present study complemented the homework time indicator by additional homework variables. It is, however, clearly necessary to expand homework research by describing the entire homework process in more detail. For instance, homework assignments differ not only in quantity (frequency, length), but also in quality. Teachers' homework attitudes may have an important impact on the quality of homework and students' homework effort. Moreover, whether students invest a little or a lot of effort in homework is determined not only by their prior knowledge, but also by motivational factors that may prove to be domain specific. Trautwein, Lüdtke, Kastens, & Köller (2006; also see Trautwein & Lüdtke, in press) recently proposed a comprehensive multilevel homework model that can be used to study these complex process in more detail (also see Warton, 2001). In their analyses, conscientiousness and homework motivation in form of expectancy and value beliefs about homework proved to be strong predictors of homework effort; perceived homework quality varied considerably between classes and predicted homework motivation and behavior. More studies are needed to test homework effects based on such a comprehensive theoretical framework.

Homework is an issue of tremendous everyday importance for students, parents, and teachers alike. Although research results do not automatically translate into educational practice (see Cooper & Valentine, 2001), given the potential impact of homework research on homework assignment and parental homework behavior, it seems especially important to apply the highest methodological standards in homework research and to disseminate results with care. Although the present investigation corroborates many previous studies in supporting the notion that the relationship between homework and achievement is a positive one, it also underlines the importance of taking a differentiated view on the matter. Homework seems to work—but homework time is not the decisive factor.

References

Allison, P. D. (2001). *Missing data*. Thousand Oaks, CA: Sage.

Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, H.-J., & Weiss, M. (2002). *Pisa 2000: Die Länder der Bundesrepublik Deutschland im Vergleich* [*PISA 2000: A comparison of student performance in the German federal states*]. Opladen, Germany: Leske + Budrich.

Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O., & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich* [*TIMSS – Mathematics and science instruction in international comparison*]. Opladen, Germany: Leske + Budrich.

Baumert, J., Roeder, P. M., Sang, F., & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen [Development of achievement and leveling of achievement differences in *Gymnasium* classes]. *Zeitschrift für Pädagogik, 32*, 639–660.

Baumert, J., Trautwein, U., & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens [School environments – Institutional conditions of learning and instruction]. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, H.-J. Tillmann, & M. Weiss (Eds.), *PISA 2000: Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (pp. 259–330). Opladen, Germany: Leske + Budrich.

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*. Chestnut Hill, MA: Boston College.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein, & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Burnett, P. C., & Fanshawe, J. P. (1997). Measuring school-related stressors in adolescents. *Journal of Youth and Adolescence, 26*, 415–428.

Carroll, J. B. (1963). A model of school learning. *Teachers' College Record, 64*, 723–733.

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*, 330–351.

Cooper, H. (1989). *Homework*. White Plains, NY: Longman.

Cooper, H. (2001). *The battle over homework* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Cooper, H., Lindsay, J. J., Nye, B., & Greathouse, S. (1998). Relationships among attitudes about homework, amount of homework assigned and completed, and student achievement. *Journal of Educational Psychology, 90*, 70–83.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*, 1–62.

Cooper, H., & Valentine, J. C. (2001). Using research to answer practical questions about homework. *Educational Psychologist, 36*, 143–153.

Cooper, H., Valentine, J. C., Nye, B., & Lindsay, J. J. (1999). Relationships between five after-school activities and academic achievement. *Journal of Educational Psychology, 91*, 369–378.

Corno, L. (1980). Individual and class level effects of parent-assisted instruction in classroom memory support strategies. *Journal of Educational Psychology, 72*, 278–292.

Corno, L. (1996). Homework is a complicated thing. *Educational Researcher, 25*, 27–30.

De Jong, R., Westerhof, K. J., & Creemers, B. P. M. (2000). Homework and student math achievement in junior high schools. *Educational Research and Evaluation, 6*, 130–157.

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*, 198–208.

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Methods for handling missing data. In J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology, Vol. 2* (pp. 87–114). New York: John Wiley & Sons.

Grolnick, W. S. (2003). *The psychology of parental control: How well-meant parenting backfires*. Mahwah, NJ: Lawrence Erlbaum Associates.

Haag, L. (1991). *Hausaufgaben am Gymnasium* [*Homework in Gymasium schools*]. Weinheim, Germany: Deutscher Studien Verlag.

Hattie, J., & Clinton, J. (2001). The assessment of teachers. *Teaching Education, 12*, 279–300.

Heller, K. A., & Perleth, Ch. (2000). *Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision (KFT 4-12+ R)* [*Cognitive abilities test for grades 4–12, Revision*]. Göttingen, Germany: Hogrefe.

Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M., & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist, 36*, 195–209.

Husén, T. (1967). *International study of achievement in mathematics. A comparison of 12 countries (Vols. I and II)*. Stockholm, Sweden: Almqvist & Wiksell.

Keith, T. Z. (1986). *Homework*. West Lafayette: Kappa Delta Pi.

Keith, T. Z., & Cool, V. A. (1992). Testing models of school learning: effects of quality of instruction, motivation, academic coursework, and homework on academic achievement. *School Psychology Quarterly, 3*, 207–226.

Keith, T. Z., Diamond-Hallam, C. D., & Goldenring Fine, J. (2004). Longitudinal effects of in-school and out-of-school homework on high school grades. *School Psychology Quarterly, 19*, 187–211.

Klieme, E., Neubrand, M., & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse [Mathematics literacy: Test conception and results]. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß

(Eds.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (pp. 139–190). Opladen, Germany: Leske + Buderich.

Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage Publications.

Martin, M. O., & Kelly, D. L. (Eds.). (1997). *Implementation and analysis. Primary and middle school years (Population 1 and Population 2). Third International Mathematics and Science Study, Vol. II*. Chesnut Hill, MA: Boston College. Technical report.

Muhlenbruck, L., Cooper, H., Nye, B., & Lindsay, J. J. (2000). Homework and achievement: explaining the different strengths of relation at the elementary and secondary school levels. *Social Psychology of Education, 3*, 295–317.

Muthén, L. K., & Muthén, B. O. (1998–2004). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.

OECD. (2001). *Knowledge and skills for life: First results from the OECD programme for international student assessment*. Paris, France: OECD.

Paschal, R. A., Weinstein, T., & Walberg, H. J. (1984). The effects of homework on learning: a quantitative synthesis. *Journal of Educational Research, 78*, 97–104.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*, (2nd ed.). Thousand Oaks, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T. (2000). *HLM 5* [*Computer program*]. Lincolnwood, IL: Scientific Software International.

Robitaille, D., & Garden, R. (1989). *The IEA study of mathematics II. Contents and outcomes of school mathematics*. Oxford, England: Pergamon Press.

Rost, J. (1996). *Lehrbuch Testtheorie – Testkonstruktion* [*A Textbook of Test Theory and Test Construction*]. Bern, Switzerland: Huber.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods, 7*, 147–177.

Schmitz, B., & Skinner, E. (1993). Perceived control, effort, and academic performance: interindividual, intraindividual, and multivariate time-series analyses. *Journal of Personality and Social Psychology, 64*, 1010–1028.

Schnyder, I., Niggli, A., Cathomas, R., Trautwein, U., & Lüdtke, O. (2006). Wer lange lernt, lernt noch lange nicht viel mehr: Korrelate der Hausaufgabenzeit im Fach Französisch und Effekte auf die Leistungsentwicklung [Low returns on time investment: correlates of time spent on French homework and effects on learning gains]. *Psychologie in Erziehung und Unterricht, 53*, 107–121.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.

Thorndike, R. L., & Hagen, E. (1993). *The cognitive abilities test: Form 5*. Itasca, IL: The Riverside Publishing Company.

Trautwein, U., & Köller, O. (2003a). The relationship between homework and achievement: still much of a mystery. *Educational Psychology Review, 15*, 115–145.

Trautwein, U., & Köller, O. (2003b). Was lange währt, wird nicht immer gut: Zur Rolle selbstregulativer Strategien bei der Hausaufgabenerledigung [Time investment does not always pay off: the role of self-regulatory strategies in homework execution]. *Zeitschrift für Pädagogische Psychologie, 17*, 199–209.

Trautwein, U., Köller, O., Schmitz, B., & Baumert, J. (2002). Do homework assignments enhance achievement? A multilevel analysis in 7th grade mathematics. *Contemporary Educational Psychology, 27*, 26–50.

Trautwein, U., & Kropf, M. (2004). Das Hausaufgabenverhalten und die Hausaufgabenmotivation von Schülern – und was ihre Eltern darüber wissen [Students' homework motivation and behavior – and their parents' knowledge of it]. *Psychologie in Erziehung und Unterricht, 51*, 285–295.

Trautwein, U., & Lüdtke, O. (in press). Students' self-reported effort and time on homework in six school subjects: between-student differences and within-student variation. *Journal of Educational Psychology*.

Trautwein, U., Lüdtke, O., Kastens, C., & Köller, O. (2006). Effort on homework in grades 5 through 9: development, motivational antecedents, and the association with effort on classwork. *Child Development, 77*, 1094–1111.

Trautwein, U., Lüdtke, O., Schnyder, I., & Niggli, A. (2006). Predicting homework effort: support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*, 438–456.

Walberg, H. J. (1991). Does homework help? *The School Community Journal, 1*, 13–15.

Walberg, H. J., & Paschal, R. A. (1995). Homework. In L. W. Anderson (Ed.), *International encyclopaedia of teaching and teacher education* (pp. 268–271). Oxford: Elsevier.

Warton, P. M. (2001). The forgotten voices in homework: views of students. *Educational Psychologist, 36*, 155–165.

Weinert, F. E., & Helmke, A. (1995). Interclassroom differences in instructional quality and interindividual differences in cognitive development. *Educational Psychologist, 30*, 15–20.

Xu, J. (2005). Purposes for doing homework reported by middle and high school students. *The Journal of Educational Research, 99*, 46–55.