# Assessment Drives Learning
## The Effect of Central Exit Exams on Curricular Knowledge and Mathematical Literacy

HENDRIK JÜRGES
KERSTIN SCHNEIDER
MARTIN SENKBEIL
CLAUS H. CARSTENSEN

# Assessment Drives Learning
## The Effect of Central Exit Exams on Curricular Knowledge and Mathematical Literacy

## Abstract

In this paper, we use data from the German PISA 2003 sample to study the effects of central exit examinations on student performance, student attitudes, and teacher behavior. Unlike earlier studies we use (i) a value-added measure to pin down the effect of central exit exams on learning in the last year before the exam and (ii) separate test scores for mathematical literacy and curriculum-based knowledge. The findings indicate that central exit exams only improve curriculum-based knowledge but do not affect mathematical literacy. Moreover, teachers in German states with central exit examinations are more active and tend to be more performance oriented. Students, although showing a better performance, are less motivated in school.

*Hendrik Jürges*
*MEA*
*University of Mannheim*
*Mannheim – Germany*
*juerges@mea.uni-mannheim.de*

*Kerstin Schneider*
*University of Wuppertal*
*Wuppertal – Germany*
*schneider@wiwi.uni-wuppertal.de*

*Martin Senkbeil*
*IPN*
*University of Kiel*
*Kiel – Germany*
*senkbeil@ipn.uni-kiel.de*

*Claus H. Carstensen*
*IPN*
*University of Kiel*
*Kiel – Germany*
*carstensen@ipn.uni-kiel.de*

May 2009

## 1. Introduction

In this paper, we present new evidence on the effects of state-mandated graduation exams or central exit examinations (CEEs) on educational outcomes in German secondary schools. The existing literature on CEE effects reaches fairly unanimous conclusions. Theoretically, centralized examinations yield positive effects on student achievement because of incentive effects on teachers and students (Costrell, 1997; Effinger and Polborn, 1999; Jürges, Richter, and Schneider, 2005). This prediction is matched by the empirical literature, although estimated effect sizes vary substantially depending on the data and identification strategy (Bishop, 1997, 1999; Jürges, Schneider, and Büchel, 2005; Wößmann, 2002). Given the beneficial effect on student achievement and provided that introducing central exit exams requires little additional monetary resources, one might be tempted to close the case. But low monetary costs do not mean that central exit exams are costless. Achievement gains might be costly for students and teachers if they put forth more effort. It is certainly of interest to find out if such costs of higher achievement exist, what they are and who eventually bears them. Do students or teachers work harder? Does teaching *quality*, a key factor in improving education, improve? (Scheerens, 2000; Scheerens and Bosker, 1997) Is learning becoming less fun? (Jürges and Schneider, 2009).

The present study makes several contributions to the empirical literature on central exit exams. First, we use longitudinal data from the PISA-I Plus study, an extension of PISA 2003 − with performance measured in ninth and tenth grade − to estimate the effect of central exit exams on student performance. Thus, in contrast to earlier studies, we can measure student achievement as value-added in the last year before the exit examination in nonacademic tracks. Second, we control for general cognitive skills of the students, as PISA-I 2003 also includes a cognitive ability test.

Our third contribution is to differentiate achievement on test items referring to basic skills (mathematical literacy) and test items that reflect the German mathematics curriculum. This difference is important, because "PISA considers student knowledge […] not in isolation but in relation to students' ability to reflect on their knowledge and experience and to apply them to real world issues." (OECD, 2003, p. 24). This type of knowledge is different from the German curriculum, which is the binding standard for schools. In contrast to the regular PISA questions, the German curriculum requires substantial technical and conceptual modeling abilities in mathematics (Blum et al., 2004). With our data, we are able to test whether

mathematical literacy and curricular knowledge are similarly affected by external standards. In that respect we can add new evidence to the "teaching to the test" question.

Fourth, we exploit the fact that the German secondary school system is characterized by a fairly rigid system of early academic tracking. Our data contain information on students and teachers in nonacademic tracks (who graduate after grade 10) and students and teachers in academic tracks (who continue schooling until grade 12 or 13). While for nonacademic tracks we measure value-added immediately before exit examinations, exit exams in academic tracks are still two or three years ahead. For students in nonacademic tracks who are in 9th and 10th grade, the incentive effects of central exit exams are therefore expected to be stronger than for their peers in the academic track. Moreover, nonacademic track students in non-CEE states graduate without taking an additional *exit* exam. A leaving certificate is awarded upon successfully completing the final grade. At the end of the academic track, however, there are exit exams in all German states, but exit exams can be administered centrally or drafted and administered locally by teachers and schools. Finally, we aim at disentangling the effect of central exit exams on student achievement by studying student and teacher attitudes and behaviour as potentially important mediators in the relationship between central exit exams and student performance.

Our findings are in line with earlier results on CEE-effects on student achievement in mathematics (Jürges, Schneider, and Büchel, 2005; Jürges and Schneider, 2009) and yield new insights into the mechanics of external standards. We show that the effects are significant only for curricular knowledge, whereas mathematical literacy appears not to be affected. Further, central exit exams appear to matter significantly only for students in nonacademic tracks. Students and teachers in 9th and 10th grade of the academic track, where students graduate after grade 12/13, are hardly affected by central exit exams. Moreover, earlier findings that students incur costs in form of being less motivated and feeling higher pressure are confirmed and teachers in states with central exit exams are more active and more achievement oriented.

The paper proceeds as follows: in Section 2, we describe relevant features of the German school system. In Section 3 we give a brief description of the PISA-I-Plus data. Sections 4 to 6 show our empirical strategy and estimation results for the effect of central exit examinations on student performance, student attitudes, and teacher behavior. Finally, we draw some conclusions in Section 7.

## 2. A primer on the German school system and recent reforms

In the context of our analysis, the German school system has two important characteristics (a comprehensive description of the German school system can be found in Jonen and Eckardt (2004)). First, it is characterized by a distinct federal structure. The 16 federal states have far-reaching autonomy in all education matters. The federal influence on education policy is rather weak and has been further weakened by the latest reform of federalism in Germany in 2006. As a result, the German school system is characterized by strong regional differences that tend to inhibit the mobility of families and teachers across German states. The diversity across states is often criticized, but the coexistence of various school systems within one country can be exploited for evaluation studies.

One important cross-state difference is the existence of central exit examinations. In some of the states, central exit examinations exist since the end of World War II. Although central exit examinations have been an ideological battlefield in Germany ever since, there was no reform until very recently. While it was always argued that students from states with central exit examinations performed much better, reliable comparisons of student achievement across states with and without central exit exams (not to speak of causal analyses) have been virtually non-existent. This was mainly due to the lack of nationwide standardized tests. In response to this lack of data, the Conference of State Education Ministries (KMK) commissioned an extension to the OECD PISA 2000 study (PISA-E), boosting sample sizes and including questions that were more specific to the German schooling system. The results of PISA-E in 2000 did indeed reveal large differences in test scores between states and in particular between states with and without exit exams. In response, all except one federal state have introduced central exit exams, however, without any reference to empirical evidence that observed differences in achievement are causally linked to central exit examinations. Further, a group of seven (since 2008: twelve) German federal states have introduced regular standardized tests of student skills at different grades in primary and secondary schools.

One reason for this recent zeal in education policy was the tremendous effect that the publication of the international PISA 2000 results had on the German public. In contrast to German self-perception, student achievement proved to be at most average in international comparison. Since then, concerns were growing about a declining quality of education in German schools (of which 95% are public) and a decreasing international competitiveness of the German labor force.

A second salient feature of the German school system is the rigid tracking in secondary schools. After primary school (usually at the age of ten), students are allocated to one of three types of secondary schools: two nonacademic tracks (basic and intermediate), and an academic track. Requirements for graduation differ across these tracks. Graduating from the academic track (in grade 12 or 13), which is equivalent to a general university-entrance certificate, requires passing an exit exam in every state. In some states, these exams are state-mandated, and in others they are designed at each individual school (subject to state-wide guidelines). A typical exit examination in academic tracks consists of four to five subjects. The choice of subjects is limited and varies from state to state, but mathematics was not mandatory in 2003 (the year that is relevant for our analysis).

**Table 1.** Overview of Central Exit Exams in Germany in 2003

|                             | Basic track | Intermediate track | Academic track |
| --------------------------- | ----------- | ------------------ | -------------- |
| Baden-Württemberg           | X           | X                  | X              |
| Bavaria                     | X           | X                  | X              |
| Mecklenburg-West Pomerania  |             | X                  | X              |
| Saarland                    | X           | X                  | X              |
| Saxony                      | X           | X                  | X              |
| Saxony-Anhalt               | X           | X                  | X              |
| Thuringia                   | X           | X                  | X              |

No Central Exit Exams in Berlin, Bremen, Hamburg, Lower Saxony, North Rhine-Westphalia, Rhineland-Palatinate, Schleswig-Holstein. Brandenburg and Hesse introduced central exit exams in 2003/2004 and are excluded from the analysis.

Leaving certificates for basic tracks are awarded upon completion of grade 9 or 10, depending on the state. Leaving certificates for intermediate tracks are awarded after grade 10. In some states, the leaving certificate requires passing a central exit examination. Table 1 describes the situation in 2003/04, the years in which the PISA-I-Plus data were collected. Hesse and Brandenburg were excluded from the analysis because they introduced CEEs in 2003/04, the year of the PISA-I study. Hence students were in grade 9 immediately before CEEs were introduced and in grade 10, when CEEs just had been introduced. It is therefore not clear how to treat the two states. Seven states had central exit examinations at the end of the intermediate track, and six had CEEs at the end of the basic track. In contrast to academic tracks, there is only limited choice of subjects in the exit exams. Written exams in German and mathematics are compulsory subjects in all central exit examinations. Science on the

other hand, is not tested in a central exam – with the exception of Saxony and to some extent Bavaria.

## 3. Data

We use data from PISA-I-Plus, an extension to PISA 2003 for Germany. Compared to the original PISA study design, this extension has several distinct characteristics. First, in contrast to the original PISA sample design that includes 15 year olds, the PISA-I-Plus baseline sample consisted exclusively of 9th-graders. Second, students had to sit two test days. On the first day, they were given the common set of international mathematics, science, and reading items. On the second day, students were given national test items. While the international test items focused on "mathematical literacy", i.e., the capacity to use mathematical skills in everyday contexts, the national items were designed to test curricular knowledge. Third, students were tested again one year later in 2004, when they were in $10^{th}$ grade. This allows us to use achievement gains as outcomes. Fourth, in addition to testing students and gathering information from student questionnaires, PISA-I-Plus has also interviewed teachers about attitudes and teaching practices.

Although our data have many advantages compared to the regular PISA data, the longitudinal design also has some limitations. Students who repeat $9^{th}$ grade, who change schools, or who leave school after grade 9 are not followed (due to the latter restriction, we excluded all students in basic tracks from our analyses). The sample thus becomes selective in terms of socio-economic background and achievement. Compared to the full sample of $9^{th}$ graders, there are more girls than boys, less students with immigration background and higher average test scores. To account for selectivity based on observable covariates, we use survey weights, so that the samples are representative for the secondary school types in Germany (Prenzel et al., 2006, pp. 46-49, 54-56).

The sample used in the analysis consists of 4,928 students, of which 2,599 students are enrolled in nonacademic track schools (with graduation after grade 10) and 2,329 students are enrolled in the academic track school (with graduation after grade 12 or 13 years). The data are described in more detail below.

## 4. Effect of exit exams on student achievement

### 4.1 Empirical strategy

In this section we report empirical estimates for the effect of exit examinations on student achievement. The basic specification is the simple education production function:

$$(1) \qquad y_{iGc} = a_{iGc} + \beta d_{iGc} + \gamma X_{iGc} + e_{iGc}$$

This equation models mathematics achievement $y$ of student $i$ in grade $G$ (grade 10) in classroom $c$ as a function of unobserved skill $a$, exam regime $d$, family (and other) background variables $X$, and a random (measurement) error term $e$. For sake of simplicity, we leave out teacher, school, and peer effects. Note that unobserved skill, $a$, is modeled as a time-varying variable, i.e., skill is explicitly allowed to evolve over time. We can think of unobserved skill as being a function of the history of all external influences on ability (family, peers, teachers, schools) and of innate learning capacity. A value-added specification of equation (1) that includes achievement in grade $G-1$ (i.e. grade 9) along with contemporaneous family characteristics and a contemporaneous exam regime dummy is specified in (2).

$$(2) \qquad y_{iGc} = \alpha_{iG} + \theta y_{i(G-1)c} + \beta d_{iGc} + \gamma X_{iGc} + e_{iGc}$$

where $\alpha_{iG}$ denotes innate learning capacity of student $i$ in grade $G$ and $\theta$ is a parameter that measures how past experiences and acquired knowledge persist into the future. Note that $y_{i(G-1)c}$ also contains past effects of the exam regime. Hence by estimating equation (2), we control for long-term effects of central exit examinations on student attainment. More importantly, including achievement in grade 9 helps to identify the causal effect of CEEs. The potential policy endogeneity of CEEs that has been discussed in detail in the literature (e.g. Jürges, Schneider, and Büchel, 2005) is taken care of in equation (2). For instance, if the decision to introduce CEEs depends on the valuation of education in a state and the attitude towards education also affects academic achievement, simple differences between CEE and non-CEE states do not estimate the causal effect of CEEs. Including achievement in grade 9,

however, controls for the influence of unobserved parental preferences for education up to grade 9. Hence our coefficient of interest, $\beta$, shows the immediate effect of central exit examinations on achievement in grade 10 controlling for past effects.

In the following, we estimate four different versions of equation (2) – with standard errors that account for stratification and clustering of the sample. First, we differentiate between students in nonacademic and in academic tracks. For each track type, we then estimate a production function for mathematical literacy and for curricular knowledge. For each individual, we have two test scores for each concept, one in grade 9 and one in grade 10. In equation (2), $y_{iGc}$ represents the test score in grade 10, and $y_{i(G-1)c}$ represents the test score in grade 9. Note that test items in grades 9 and 10 were not the same. Each test was an independent assessment of mathematical literacy and curricular knowledge – standardized to mean 50 and standard deviation 10. Hence we do not measure absolute but relative achievement gains. Additional information on the test and the test items can be found in Blum et al. (2004). Descriptive statistics for test scores, by exam and track type, are shown in Table 2. Two results stand out. First, academic track students have on average much higher PISA test scores in all tests, independent of the exam type. The difference amounts to one standard deviation in the overall test score distribution. Second, students in CEE states have higher average PISA test scores, independent of the track type. The difference is between 0.1 and 0.2 standard deviations.

For innate learning capacity $\alpha_{iG}$, we use two proxy variables: an index for cognitive skills that is based on test items on figure analogies, as part of a non-verbal IQ-test (German revised adaptation of the Thorndike and Hagan-cognitive abilities test (Heller and Perleth, 2000)), and a dummy variable that reflects whether a student already had to repeat a grade. For general cognitive skills, Table 2 shows a large difference between nonacademic and academic track students – independent of the exam type and repeating classes is more common in nonacademic than in academic tracks. However, the general cognitive skills score difference between the two types of tracks is smaller than the test score difference. An explanation for this finding is that the tracked school system in Germany reinforces innate ability differences between students. Another important finding is that general skills do not differ significantly between students in CEE and non CEE states.

As indicators of the social background we use the international socio-economic index of the highest parental occupational status (HISEI) (Ganzeboom *et al.,* 1992, OECD, 2005) and a dummy variable for immigration background that has the value of one if at least one parent

was not born in Germany. Table 2 shows that students in academic track schools have considerably higher values of the socio-economic index. Further, 12 percent of the academic-track students in CEE states have a migration background compared to 14 percent in non-CEE-states. This difference is more pronounced for students in the nonacademic track. While in CEE states only 10 percent have a migration background, in non-CEE states the percentage amounts to 21 percent.

**Table 2.** Description of student sample

| | Non-academic students | | | | Academic students | | | |
|---|---|---|---|---|---|---|---|---|
| | CEE (N = 1453) | | No CEE (N = 1146) | | CEE (N = 1079) | | No CEE (N = 1250) | |
| | Mean | (S.E.) | mean | (S.E.) | mean | (S.E.) | mean | (S.E.) |
| HISEI | 47.83 | (0.58) | 49.63 | (0.62) | 59.36 | (0.69) | 59.61 | (0.69) |
| Cognitive skills | 46.87 | (0.40) | 47.22 | (0.47) | 54.20 | (0.55) | 53.96 | (0.38) |
| Immigration background | 0.10 | (0.01) | 0.21 | (0.02) | 0.12 | (0.02) | 0.14 | (0.01) |
| Repeat class | 0.21 | (0.02) | 0.22 | (0.02) | 0.07 | (0.01) | 0.10 | (0.01) |
| Mathematical literacy test grade 9 | 46.24 | (0.44) | 44.55 | (0.44) | 56.39 | (0.47) | 54.96 | (0.39) |
| Mathematical literacy test grade 10 | 46.40 | (0.43) | 45.47 | (0.47) | 55.80 | (0.48) | 54.34 | (0.44) |
| Mathematical curricular test grade 9 | 46.11 | (0.46) | 44.48 | (0.51) | 56.88 | (0.50) | 54.75 | (0.47) |
| Mathematical curricular test grade 10 | 46.80 | (0.45) | 43.74 | (0.49) | 56.53 | (0.50) | 54.83 | (0.44) |

CEE: central exit exams

S.E.: standard error of the mean

*4.2 Results*

Table 3 shows our regression results. The first two models show regression results with mathematical literacy in grade 10 as dependent variable. In model (1) we restrict the sample to students in nonacademic tracks, whereas model (2) includes only students in academic track schools. The main finding is that there are only small and insignificant coefficients of central exit examinations on mathematical literacy in grade 10. In fact, in model (1) the point estimate for students who graduate after grade 10 from nonacademic tracks is even negative. Put differently, there is no evidence that central exit examinations in mathematics improve mathematical literacy.

With respect to our control variables we briefly note that, as expected, general cognitive ability is a highly significant predictor of mathematical literacy. Also, students who did not repeat a class until grade 9 have significantly higher achievement gains. The strongest predictor for mathematical literacy in grade 10 is mathematical literacy in grade 9. Students

with parents of higher socio-economic status have slightly higher test scores in mathematical literacy but the coefficients are not significant. Similarly, immigration background has a small but insignificant effect. Most likely, the effects of the social background are absorbed by general cognitive ability and by test scores measured in grade 9. The gap in mathematical literacy is not further widened in grade 10 (e.g., Ai, 2002; Scott et al., 1995).

In models (3) and (4), the dependent variable is the score on the curricular knowledge test. In contrast to our results for mathematical literacy, we now find a highly significant positive effect of central exit exams on achievement gains in the curriculum-based knowledge test. However, this finding is constrained to students who graduate after grade 10, i.e., for students in the nonacademic track. For academic track students the CEE effect remains insignificant and fairly weak. There are two possible explanations for this lack of effect. First, academic track students have more time until graduation which weakens any incentive effects if students discount the future. Second, even if exams at the end of the academic track school are not administered centrally, there is an exit exam. These exams are subject to the approval by the supervisory authority, so that teachers are not entirely free in setting up the test problems. Hence there might be some effect of the non-central exit exam as well.

**Table 3.** Value-added regression of student achievement

| | Mathematical literacy test | | | | Mathematical curricular test | | | |
| | Nonacademic track | | Academic track | | Nonacademic track | | Academic track | |
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | |
| | b | (S.E.) | b | (S.E.) | b | (S.E.) | b | (S.E.) |
|---|---|---|---|---|---|---|---|---|
| Mathematical literacy grade 9 | 0.62*** | (0.05) | 0.62*** | (0.06) | 0.69*** | (0.05) | 0.68*** | (0.05) |
| Cognitive skills | 0.20*** | (0.03) | 0.16*** | (0.03) | 0.12*** | (0.03) | 0.08*** | (0.03) |
| No repeat class | 0.83** | (0.33) | 1.31** | (0.56) | 0.69* | (0.38) | 0.91 | (0.57) |
| HISEI | 0.01 | (0.01) | 0.02 | (0.01) | 0.02 | (0.01) | 0.03*** | (0.01) |
| No immigration background | 0.24 | (0.39) | 0.59 | (0.55) | 0.18 | (0.39) | 1.17*** | (0.42) |
| Central exit exams | -0.13 | (0.36) | 0.29 | (0.41) | 1.95*** | (0.33) | 0.22 | (0.41) |
| $R^2$ | 0.57 | | 0.50 | | 0.61 | | 0.57 | |
| N | 1453 | | 1146 | | 1079 | | 1250 | |

S.E.: standard error of the mean

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

In terms of control variables, we find again that general cognitive ability and test scores in grade 9 are the strongest predictors for test scores in grade 10. However, compared to the results obtained for mathematical literacy in (1) and (2), the relative importance of the grade 9 test score grows and the influence of general cognitive skills decreases. This is a plausible result, because we expect general cognitive ability to be a better predictor of general skills (mathematical literacy) than of specific skills (curricular knowledge).

For academic track students, family background is statistically significant in explaining achievement gains in curricular knowledge. For nonacademic track students family background remains insignificant. This is noteworthy because students in academic track schools tend to come from families with comparably high socioeconomic status and only a small proportion of students with a migration background attend the academic track. Hence one might expect the differences in socio-economic background to be more relevant for nonacademic track students. This is not the case. The socioeconomic gap in academic achievement continues to widen for academic track schools, possibly because parents with a high socio-economic status can either help their children themselves or organize and finance coaching to prepare for the exams. By the same token, having a migration background is detrimental for progress in curriculum-based knowledge for academic track students.

In summary, our analysis in this section suggests that achievement gains in mathematical literacy between grades 9 and 10 are not larger in states with central exit exams. Gains in curricular knowledge, however, are significantly larger in states with central exit exams, but only for students who finish school after grade 10. This insight is important for our understanding of the mechanics of central exams. If students are tested centrally and students and teachers know that the command of curricular knowledge is tested, curricular knowledge increases. Mathematical literacy (as a broader concept) is not in the focus of the curriculum and hence the exit exam. Thus there are no strong incentives to improve on the students' literacy skills. This is consistent with the education literature: Teachers teach the curriculum if the curriculum is tested (Au, 2007; Jacob, 2005).

We interpret our result as an argument in favor of CEEs, as the results show that students have a better knowledge of the material they are tested on, which is defined in the curriculum. In Germany the curriculum is binding on the state level and teachers cannot adjust the number of mathematics lessons and divert resources from other subjects to mathematics. The number of mathematics lessons is fixed and mathematics teachers rarely teach a second subject in the same class. Hence with our data, we get an estimate of the net CEE effect. The analysis also shows

that one has to be precise about the contents of the curriculum or the abilities that students are expected to have at the end of schooling. Mathematical literacy is not automatically improved if the curriculum requires fairly abstract technical and conceptual modeling abilities. To increase mathematical literacy, one might think about means on how to include literacy in the exit exams and also in the curriculum.

## 5. Effect of exit exams on teacher attitudes and behavior

*5.1 Empirical Strategy*

Having found a positive effect of central exit exams in nonacademic track schools for curriculum-based knowledge, we now try to identify possible explanations for the CEE-effect (Jürges, Richter, and Schneider, 2005; Jürges and Schneider, 2009). In particular, it might be argued that CEEs enhance the quality of teaching and therefore raise student achievement (Scheerens, 2000; Scheerens and Bosker, 1997). Teaching quality, however, is not directly measurable, but PISA-I-Plus contains rich information on teacher attitudes and behavior that was collected in a supplementary teacher questionnaire in 2003. In this section, we analyze differences in teacher attitudes and behavior between states with and without central exit exams. Note, however, that the available information can only serve as more or less useful proxies for the quality of teaching (Seidel and Shavelson, 2007). Hence our results ought to be interpreted with some caution.

We restrict our analytical sample to mathematics teachers who taught the tested students in grades 9 and 10. This ensures that classes have been taught by the same teacher between the two tests. Our sample restriction leaves us with a sample of 134 teachers or classes, respectively. 69 of these classes are in states with central exit exams and the remaining 65 classes are in states without central exit exams. We also compare teacher attitudes and behavior of teachers who did not teach a centrally tested subject. For this purpose, we selected 56 science teachers, who reported not to teach German, mathematics or English. Using this control group allows to strengthen our results. If central exit exams affect teaching, the effect should exist only for teachers, who teach subjects that are tested centrally, i.e. mathematics teachers. Students in the nonacademic track schools are typically not tested centrally in science. Hence we expect to find no differences in science teacher attitudes and behavior between states with and without exit exams. The only exception for the nonacademic track is Saxony, where science is also tested

centrally. Therefore science teachers from Saxony are not included in the control group. In Bavaria, some students are tested in physics but not in other sciences. We decided to include the Bavarian teachers in the sample because it is not clear which teachers are affected. Moreover, since we expect science teachers not to be affected by CEEs as opposed to mathematics teachers, including some teachers who are affected by CEEs in the control group will bias our results towards zero and against finding a CEE effect. For teachers in the academic track, where both science and mathematics might be a tested subject in the exit exam, no differences between the types of teachers are expected.

Overall, we used 76 items from the teacher questionnaire to measure attitudes and behavior of mathematics and science teachers. From these items, we constructed indices for six different dimensions: performance orientation, teacher cooperation, effective time use, fostering of learning and disciplinary climate, parent involvement, and use of evaluation methods. Table 4 contains variable definitions and descriptive statistics by type of exit examinations and school type.

The basic specification for our regression is:

$$(3) \qquad y_c = \beta d_c + \gamma \overline{X}_c + \varepsilon_c$$

where $y_c$ is the (z-standardized) attitudes score for teacher $c$, $d_c$ is a dummy variable for central exit exams, and $\overline{X}_c$ denotes average general cognitive ability, socio-economic status, immigration background of the students in the class, and the size of the school of teacher $c$.

Using a subset of 15 items, we also identified three types of teachers (active, performance oriented, passive) by latent class analysis (for details see Senkbeil, 2006). This classification allows us to study possible effects of exit exams on the six dimensions of teacher attitudes and behavior simultaneously. "Active" teachers (41 percent of the sample) have high values in all of the dimensions mentioned above. "Performance oriented" teachers (also 41 percent of the sample) stress performance and effective time use, but are less interested in cooperation and evaluation. "Passive teachers" (18 percent) have low scores on all of the six dimensions shown above. We test for differences in teacher types across different types of exams by cross-tabulation and using the chi-squared test.

**Table 4.** Description of teacher variables

| | # of items | Cronbach's alpha | Example item | Non-academic students | | | | Academic students | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | CEE | | Non CEE | | CEE | | Non CEE | |
| | | | | Mean | (S.E.) | mean | (S.E.) | mean | (S.E.) | mean | (S.E.) |
| **Mathematics teachers** | | | | | | | | | | | |
| Parent involvement | 5 | 0.69 | I want parents to talk to the teachers as often as possible. | 0.05 | (0.17) | 0.04 | (0.18) | -0.08 | (0.17) | -0.05 | (0.18) |
| Use of evaluation methods | 20 | 0.88 | There are defined standards for each class level at our school. | 0.02 | (0.18) | -0.05 | (0.15) | -0.13 | (0.15) | 0.15 | (0.20) |
| Fostering of learning / disciplinary climate | 9 | 0.82 | We monitor the pupils' performance consistently. | 0.16 | (0.14) | -0.38 | (0.16) | 0.46 | (0.20) | -0.17 | (0.18) |
| Teacher cooperation | 20 | 0.90 | How many times do you meet with mathematics teachers to discuss or plan content of the curriculum or teaching methods? | 0.25 | (0.17) | -0.14 | (0.14) | -0.22 | (0.17) | 0.02 | (0.21) |
| Performance orientation | 12 | 0.82 | High demands are made on the pupils at our school. | 0.28 | (0.14) | -0.38 | (0.16) | 0.19 | (0.17) | -0.12 | (0.21) |
| Effective time use | 10 | 0.71 | It is important to us to free up as much working time as possible. | 0.12 | (0.15) | -0.09 | (0.18) | 0.10 | (0.18) | -0.15 | (0.19) |
| | | | | | | | | | | | |
| **Science teachers** | | | | | | | | | | | |
| Parent involvement | 5 | 0.69 | I want parents to talk to the teachers as often as possible. | 0.12 | (0.25) | -0.13 | (0.24) | 0.46 | (0.32) | -0.21 | (0.26) |
| Use of evaluation methods | 20 | 0.88 | There are defined standards for each class level at our school. | 0.02 | (0.31) | -0.01 | (0.21) | -0.03 | (0.41) | 0.01 | (0.20) |
| Fostering of learning / disciplinary climate | 9 | 0.82 | We try to understand the problems of the pupils. | -0.21 | (0.29) | 0.19 | (0.21) | 0.44 | (0.40) | -0.29 | (0.21) |
| Teacher cooperation | 20 | 0.90 | How many times do you meet with other teachers to discuss or plan content of the curriculum or teaching methods? | 0.14 | (0.26) | 0.08 | (0.24) | -0.32 | (0.46) | -0.21 | (0.19) |
| Performance orientation | 12 | 0.82 | High demands are made on the pupils at our school. | -0.05 | (0.23) | -0.08 | (0.23) | 0.32 | (0.52) | 0.06 | (0.27) |
| Effective time use | 5 | 0.71 | It is important to us to free up as much working time as possible. | -0.10 | (0.22) | 0.13 | (0.27) | -0.06 | (0.40) | -0.01 | (0.28) |

Table 5 shows the estimation results for the effect of exit examinations separately for each of the six teacher characteristics. In non-academic track, we find that mathematics teachers in CEE states generally rely more heavily on discipline and achievement. With respect to the other dimensions, there are no significant differences between CEE and non CEE-states. As expected, science teachers in the nonacademic track show no significant differences in attitude and behavior. Academic track teachers in CEE states are more engaged in fostering the learning and disciplinary climate. This difference is significant for mathematics and science teachers. In addition science teachers in CEE-states appear to have a more effective time use. Regarding performance orientation we find no significant differences.

In Table 6 we show the distribution of teacher "types" by exam regime, separately for mathematics and science teachers. Mathematics teachers in nonacademic tracks who teach in states with central exit exams are significantly more often "active" or "performance oriented", and less often "passive" than mathematics teachers in states without central exit exams. In contrast, mathematics teachers in academic tracks who teach in states with central exit exams are about equally likely to be either "active", "performance oriented" or "passive". Also, science teachers in both academic and non-academic tracks show no significant differences across exam types.

To summarize the results in this section, we find some evidence for differences between teachers that are exposed to central exit exams and teachers who are not. The main difference identified with our data is the promotion of positive learning and disciplinary climate. Moreover, the proportion of active or performance oriented teachers is higher in nonacademic track schools in CEE states. Teachers in academic track schools are more similar with respect to our measured dimensions.

**Table 5.** Differences in teacher attitudes and behavior[a]

| | Mathematics teachers Mean CEE minus non-CEE | | Science teachers Mean CEE minus non-CEE | |
|---|---|---|---|---|
| | d | t | d | t |
| Non-academic track | | | | |
| Parent involvement | -0.12 | 0.446 | -0.44 | 1.090 |
| Use of evaluation methods | -0.10 | 0.382 | -0.34 | 0.789 |
| Fostering of learning/disciplinary climate | 0.41* | 1.674 | -0.66 | 1.529 |
| Teacher cooperation | 0.17 | 0.740 | -0.29 | 0.759 |
| Performance orientation | 0.44* | 1.887 | 0.21 | 0.599 |
| Effective time use | 0.20 | 0.776 | 0.09 | 0.200 |
| | | | | |
| Academic track | | | | |
| Parent involvement | -0.03 | 0.316 | 1.30 | 1.643 |
| Use of evaluation methods | -0.32 | 1.132 | 0.65 | 0.716 |
| Fostering of learning/disciplinary climate | 0.68** | 2.368 | 1.28** | 2.406 |
| Teacher cooperation | -0.40 | 1.362 | 0.68 | 1.010 |
| Performance orientation | 0.29 | 0.982 | -0.67 | 0.584 |
| Effective time use | 0.19 | 0.619 | 1.51* | 1.868 |

[a] Controlling for cognitive ability, socio-economic status and immigration background of the students, sex and age of the teachers and school size.

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

**Table 6.** Distribution of teacher types

| | Mathematics teachers | | | | Science teachers | | | |
|---|---|---|---|---|---|---|---|---|
| | Active teachers | Performance oriented teachers | Passive teachers | $\chi^2$ | Active teachers | Performance oriented teachers | Passive teachers | $\chi^2$ |
| Non-academic track | | | | | | | | |
| CEE | 54.76 | 35.71 | 9.52 | 12.27*** | 33.33 | 44.44 | 22.22 | 1.20 |
| Non CEE | 19.44 | 47.22 | 33.33 | | 38.10 | 52.38 | 9.52 | |
| | | | | | | | | |
| Academic track | | | | | | | | |
| CEE | 29.63 | 62.96 | 7.41 | 1.86 | 33.33 | 50.00 | 16.67 | 0.58 |
| Non CEE | 44.83 | 44.83 | 10.34 | | 18.18 | 54.55 | 27.27 | |

* p < 0.10

** p< 0.05

*** p < 0.01

## 6. Effect of exit exams on student attitudes

### 6.1 Empirical Strategy

In this section, we use data from the student questionnaires to construct four indices of student attitudes in (grade 9) with respect to the learning climate in mathematics: *Emotions* (anger, anxiety, boredom, despair, achievement pressure), *work habits* (effort, completion of homework, attention in lessons), *motivation* (instrumental motivation, enjoyment, performance orientation), and *self esteem* (self efficacy, self concept). Table 7 contains variable definitions and descriptive statistics by type of exit examinations and school type. Note that values for each item were z-standardized. Positive values correspond to higher than average agreement, negative values correspond to less than average agreement.

Our basic regression model is:

$$(4) \qquad r_{iGc} = a_{iGc} + \beta d_{iGc} + \gamma X_{iGc} + e_{iGc},$$

where attitudes *r* reported by student *i* in grade *G* in class *c* are modeled as a function of general cognitive ability *a*, the exam regime *d* and background variables *X* (socioeconomic and migration background). Again, cognitive ability is approximated by the cognitive skills score and a dummy variable that indicates whether the student has repeated a class. Ability and background are included as control variables they might not only affect achievement but also motivation and the self perception with respect to mathematics.

### 6.2 Results

Table 8 summarizes the results. It turns out that the differences in student traits and attitudes are fairly small in magnitude. Nevertheless, students in states with central exit exams show significantly more often negative emotions like anger, anxiety, achievement pressure and despair and the self-concept in mathematical competence is relatively weak. This is in particular true for students who are at the end of the nonacademic track, but differences are also significant for students in the academic track.

**Table 7.** Sample description of students` attitudes

| | # of items | Alpha | Example item | Non-academic students | | | | Academic students | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | CEE | | Non CEE | | CEE | | Non CEE | |
| | | | | mean | (S.E.) | mean | (S.E.) | mean | (S.E.) | mean | (S.E.) |
| *Emotions* | | | | | | | | | | | |
| Anger | 5 | 0.82 | I am annoyed that mathematics is so hard. | 0.14 | (0.03) | -0.12 | (0.05) | 0.06 | (0.05) | -0.13 | (0.05) |
| Anxiety | 9 | 0.89 | Right before mathematics exams I am very nervous. | 0.15 | (0.04) | -0.08 | (0.04) | 0.02 | (0.04) | -0.16 | (0.04) |
| Boredom | 5 | 0.86 | I am very bored during mathematics classes. | 0.02 | (0.04) | -0.08 | (0.04) | 0.08 | (0.04) | -0.03 | (0.06) |
| Despair | 5 | 0.88 | During mathematics exams the idea to give up is very appealing. | 0.14 | (0.03) | -0.10 | (0.04) | 0.01 | (0.04) | -0.11 | (0.04) |
| Achievement pressure | 6 | 0.79 | It is very important to my parents for me to be good at mathematics. | 0.18 | (0.04) | 0.05 | (0.06) | -0.11 | (0.06) | -0.19 | (0.05) |
| *Work habits* | | | | | | | | | | | |
| Effort | 5 | 0.80 | I put a lot of effort into mathematics to understand everything. | 0.01 | (0.03) | 0.05 | (0.03) | -0.05 | (0.04) | -0.03 | (0.04) |
| Completion of homework | 10 | 0.72 | I always try to solve all the problems in my mathematics homework. | -0.01 | (0.04) | 0.14 | (0.03) | -0.17 | (0.03) | 0.02 | (0.04) |
| Attention in lessons | 3 | 0.76 | Even if the subject in mathematics is very difficult I put all my effort into understanding it. | -0.05 | (0.04) | 0.13 | (0.03) | -0.13 | (0.04) | 0.06 | (0.05) |
| *Motivation* | | | | | | | | | | | |
| Instrumental motivation | 4 | 0.82 | I see that mathematics will be very expedient in my future job, so I apply myself to it. | 0.00 | (0.03) | 0.08 | (0.04) | -0.11 | (0.05) | 0.01 | (0.04) |
| Enjoyment | 6 | 0.92 | Mathematics is fun. | -0.04 | (0.05) | 0.17 | (0.04) | -0.16 | (0.05) | 0.03 | (0.06) |
| Performance orientation | 3 | 0.62 | I want my grades to be good so I make an effort to achieve that. | 0.05 | (0.04) | 0.03 | (0.03) | -0.05 | (0.03) | -0.05 | (0.04) |
| *Self Esteem* | | | | | | | | | | | |
| Self efficacy | 8 | 0.81 | I can calculate how much cheaper a TV set is at a 30% discount. | -0.24 | (0.03) | -0.20 | (0.04) | 0.25 | (0.05) | 0.31 | (0.04) |
| Self concept | 5 | 0.92 | I keep up easily in mathematics. | -0.09 | (0.04) | 0.13 | (0.04) | -0.11 | (0.04) | 0.10 | (0.04) |

CEE: central exit exams

S.E.: standard error of the mean

**Table 8.** Differences in student attitudes by exam type[a]

| | Non-academic students | | Academic students | |
| | Mean CEE minus Non-CEE | | Mean CEE minus Non-CEE | |
| | d | t | d | T |
|---|---|---|---|---|
| *Emotions* | | | | |
| Anger | 0.24*** | 3.899 | 0.21*** | 2.945 |
| Anxiety | 0.20*** | 3.796 | 0.17** | 2.620 |
| Boredom | 0.09 | 1.361 | 0.10 | 1.483 |
| Despair | 0.20*** | 3.512 | 0.13** | 2.114 |
| Achievement pressure | 0.18** | 2.284 | 0.12 | 1.437 |
| *Work habits* | | | | |
| Effort | -0.05 | -1.086 | -0.03 | -0.598 |
| Completion of homework | -0.15*** | -2.857 | -0.22*** | -4.246 |
| Attention in lessons | -0.14*** | -2.801 | -0.20*** | -3.025 |
| *Motivation* | | | | |
| Instrumental motivation | -0.06 | -1.086 | -0.11** | -2.023 |
| Enjoyment | -0.18*** | -2.724 | -0.18** | -2.253 |
| Performance orientation | 0.02 | 0.435 | -0.01 | -0.128 |
| *Self Esteem* | | | | |
| Self efficacy | -0.01 | -0.194 | -0.04 | -0.668 |
| Self concept | -0.18*** | -3.127 | -0.20*** | -3.454 |

[a] Controlling for the socioeconomic background, migration background, the cognitive skills of the student and whether the student has repeated a class

* $p < 0.10$

** $p < 0.05$

*** $p < 0.01$

Differences in motivational orientation (for instance joy, homework) are also significant. It is striking that differences between CEE and non-CEE states are primarily significant when they relate to negative student attitudes. The expected positive effects from more effective and achievement oriented teaching are at least not perceived by the students. Moreover, differences between CEE and non-CEE states prevail as well for academic track students. Recall that neither achievement nor teacher characteristics differed between academic track schools.

Our results are in line with recent studies in which standardised tests (or high-stake tests) are understood as a highly controlling extrinsic motivation strategy. Although this strategy was often found to have the desired effect (better performance), it often also had a number of undesirable side effects, e.g. loss of intrinsic subject motivation, increased test anxiety, increased pressure to perform, lower self-efficacy (Abrams et al., 2003; Ryan, Ryan, Arbuthnot and Samuels, 2007; Ryan and Sapp, 2005). The comparatively small effect size

may indicate that central exit examinations do not have a homogenous effect on all students: while they provide an (external) incentive to achieve their maximum performance in a test situation for some students, other students (in particular, those who do not have much self-confidence in their own abilities) may see them as a threat or a hurdle which they cannot manage and this then leads to the undesirable side effects on motivation mentioned above (cf., e.g. Abrams et al., 2003; Kellaghan et al., 1996; Ryan et al., 2007). It is surprising, however, that these negative effects are also found among students of academic tracks. The undesirable effects on students' motivation possibly occur independent of teacher characteristics and behaviour.

## 7. Summary and Conclusion

Using data from the German PISA-I-Plus study of 2003, we have analyzed the effect of central exit exams in mathematics on students' mathematics performance, students' attitudes towards mathematics, and mathematics teachers' attitudes and behavior. The PISA-I-Plus data are superior to data used in earlier studies for two reasons. First, they contain repeated measures of performance in grades 9 and 10 (when exams take place), so that we could measure actual gains in achievement. Second, the data also allow differentiating between achievement on a mathematical literacy test and a test of curricular knowledge. While central exit exams significantly increase curricular knowledge of students, we found no significant effect on mathematical literacy – the type of mathematical knowledge that is regarded to be relevant in everyday situations. This paper thus qualifies earlier findings on the central exam effect in Germany in an important way: measured (incentive) effects of central exit examination appear to be larger when outcome measures are more in line with what is actually tested in central exit examinations.

Only few previous studies have attempted to shed light on the question what is driving the positive achievement gains of central exit exams. Theoretically, teachers as well as students should react to central exit exams by increasing effort. Empirical results are less clear. For instance, using student-provided data on teaching styles from TIMSS 1995, Jürges and Schneider (2009) find no differences in teaching practices other than teachers in CEE-states giving and checking homework more often. However, students might not be the best judges of the quality of schooling (Kunter and Baumert, 2006). In contrast to the international PISA data, PISA-I-Plus 2003 does contain detailed information on teaching practices and

20

teacher attitudes provided by the teachers themselves. Using these data, we find evidence that central exit exams have an (albeit limited) effect on teachers. Teachers tend to be more active and performance oriented when their students have to pass a central exit exam. Using student self-ratings on behavior, motivation, and attitudes towards learning mathematics, Jürges and Schneider (2009) find differences in general student motivation. There appears to be a downside to CEEs because students in CEE-states do like mathematics less. They find it less easy and more boring than those in non-CEE states. The present study confirms findings that central exit exams are associated with negative student attitudes: students in CEE-states are generally more anxious, feel higher achievement pressure, and are actually less motivated to learn.

In a further step of analysis – not reported in detail – we also analyzed whether the achievement gains we saw earlier can be explained by differences in teacher attitudes and behavior across states with and without central exit exams. In other words, we tried to figure out whether effects of central exit exams on teachers carry over to higher achievement. However, we were not able to find strong support of the hypothesis that teacher variables have this mediating effect. One problem might be that the answers to the subjective teacher questions are hard to compare across states. Since there is only limited mobility of teachers across states, the (self)-assessment of teacher characteristics depends on the characteristics of the peer group of teachers subject to central exit exams. If for instance performance orientation in CEE states is, for whatever reason, generally higher in CEE-states, the answers might not be comparable across states. Another problem could be that the available information does not capture the relevant aspects of teacher quality. Also, CEEs might simply not improve achievement because of higher teaching quality, but because of more student effort. In that case, students alone bear the cost of higher achievement. Even given the (rich) data in PISA-I 2003, we are still not able to give a final answer to the question why CEEs increase performance.

One important result of this paper is that differences in achievement due to central exit exams are typically found only in nonacademic tracks. Students and teachers in academic track schools appear to be less affected. There are various possible explanations for this result. One explanation is the difference in years until graduation that is reducing the CEE-effect. Another explanation is that (centrally approved) exit exams exist for students at academic track schools in all German states, even though they are not centrally drafted and administered. Thus, the curriculum is more binding at academic tracks schools even without

central exit exams. The most degrees of freedom with respect to the academic level required for graduation exist in lower level (basic track) secondary schools, where no formal exit exam is required, although there are of course exams during the final grade. Overall, the results of this study support the monitoring of student achievement by standardized tests at all levels of schooling and all levels of tracks, as that enhances performance.

The results also suggest that if achievement is monitored e.g. by central exit exams, teachers do in fact teach the curriculum, whereas mathematical literacy (as a broader concept) does not automatically improve. Notably, these differences between curricular and literacy tests are in accordance with recent findings from the US (e.g. Amrein and Berliner, 2002; Klein et al., 2000; Neil and Gaylor, 2001), where state-mandated exams were found to increase performance measured in terms of the curriculum but not necessarily in terms of a deeper understanding of mathematical concepts that enables students to actually use mathematics in everyday life. To improve the quality of education by central exit exams thus may require including strategies of knowledge transfer into the curriculum.

# References

Abrams, L. M., Pedulla, J. J., Madaus, G. F., 2003. Views from the classroom: Teachers` opinions of statewide testing programs. Theory into Practice 42(1), 18-29.

Ai, X., 2002. Gender differences in growth in mathematics achievement: Three-level longitudinal and multilevel analysis of individual, home, and school influences. Mathematical Thinking and Learning 4, 1-22.

Amrein, A., Berliner, D. C., 2002. High-stakes testing, uncertainty, and student learning. Education Policy Analysis Archives, 10 (18). Retrieved April 23, 2009 from http://epaa.asu.edu/epaa/v10n18/.

Au, W., 2007. High-stakes testing and curricular control: A qualitative metasynthesis. Educational Researcher 36 (5), 258-267.

Bishop, J. H., 1997. The effect of national standards and curriculum-based exams on achievement. American Economic Review 87, 260-264.

Bishop, J. H., 1999. Are national exit examinations important for educational efficiency? Swedish Economic Policy Review 6, 349-401.

Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ulfig, F., Carstensen, C. H., 2004. Mathematische Kompetenz [Mathematical competence], in: PISA-Konsortium Deutschland (Eds.), PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Waxmann, Münster, pp. 47-92.

Costrell, R. M., 1997. Can educational standards raise welfare? Journal of Public Economics 65, 271-293.

De Fraine, B., van Damme, J., Onghena, P., 2002. Accountability of schools and teachers: What should be taken into account? European Educational Research Journal 1, 403-428.

Effinger, M. R., Polborn, M. K., 1999. A model of vertically differentiated education. Journal of Economics 69, 53-69.

Ganzeboom, H. B. G., de Graaf, P., Treiman, D. J., 1992. A standard international socio-economic index of occupational status. Social Science Research 21, 1-56.

Heller, K. A., Perleth, C., 2000. Kognitiver Fähigkeitstest (KFT 4-12+R). (German revised adaptation of the Cognitive Abilities Test; Thorndike and Hagan, 1954). Hogrefe, Göttingen.

Jacob, B. A., 2005. Accountability, Incentives And Behavior: The Impact Of High-Stakes Testing In The Chicago Public Schools, Journal of Public Economics 89, 761-796.

Jonen, G., Eckardt, T., 2006. The Education System in the Federal Republic of Germany 2004. Secretariat of the Standing Conference of the Ministers of Education (KMK). http://www.kmk.org/ dossier/dossier_en_ebook.pdf [2007, Jan 12].

Jürges, H., Richter, W.F., Schneider, K., 2005. Teacher quality and incentives, Theoretical and empirical effects of standards on teacher quality. FinanzArchiv 61 (3), 298-326.

Jürges, H., Schneider, K., Büchel, F., 2005. The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany. Journal of the European Economic Association 3 (5), 1134-1155.

Jürges, H., Schneider, K., 2009. Central exit examinations increase performance...but take the fun out of mathematics. Journal of Population Economics (in press).

Kellaghan, T., Madaus, G. F., Raczek, A., 1996. The use of external examinations to improve student motivation. Washington, DC: American Educational Research Association.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., 2000. What do test scores in Texas tell us? Education Policy Analysis Archives, 8 (49). Retrieved April 23, 2009 from http://epaa.asu.edu/epaa/v10n18/.

Kunter, M., Baumert, J., 2006. Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. Learning Environmental Research 9, 231-251.

23

Neil, M., Gaylor, L., 2001. Do high-stakes graduation tests improve learning outcomes? In G. Orfield and M. L. Kornhaber (Eds.), Raising standards or raising barriers? Inequality and high-stakes testing in public education. The Century Foundation Press, New York, pp. 107-126.

OECD, 2003. The PISA 2003 assessment framework. Mathematics, reading, science and problem solving knowledge and skills. OECD, Paris.

OECD, 2005. PISA 2003. Technical report. OECD, Paris.

Prenzel, M., Carstensen, C. H., Schöps, K., Maurischat, C., 2006. Die Anlage des Längsschnitts bei PISA 2003 [The design of the longitudinal study of PISA 2003]. In PISA-Konsortium Deutschland (Hrsg.), PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres. Waxmann, Münster, pp. 29-62.

Ryan, K. E., Ryan, A. M., Arbuthnot, K., Samuels, M., 2007. Students` motivation for standardized math exams. Educational Researcher 36(1), 5-13.

Ryan, R. M., Sapp, A., 2005. Considering the impact of test-based reforms: A self-determination theory perspective on hight stakes testing and student motivation and performance. Unterrichtswissenschaft 33, 143-159.

Scheerens, J., 2000. Improving school effectiveness. UNESCO: International Institute for Educational Planning, Paris.

Scheerens, J., Bosker, R., 1997. The foundations of educational effectiveness. Elsevier Science Ltd, Oxford.

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, 454-499.

Senkbeil, M., 2006. Die Bedeutung schulischer Faktoren für die Kompetenzentwicklung in Mathematik und in den Naturwissenschaften [The importance of school factors for skill acquisition in mathematics and the natural sciences]. In PISA-Konsortium Deutschland (Hrsg.), PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres. Waxmann, Münster, pp. 277-308.

Scott, L. A. et al., 1995. Two years later: Cognitive gains and school transitions of NELS:88 eighth graders. National Center for Education Statistics, Washington.

Wößmann, L., 2002. Central Exams Improve Educational Performance: International Evidence. Kiel Discussion Papers 397, Kiel Institute for World Economics.

# CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

_____

2651 Augusto de la Torre, María Soledad Martínez Pería and Sergio L. Schmukler, Drivers and Obstacles to Banking SMEs: The Role of Competition and the Institutional Framework, May 2009

2652 Tobias Lindhe and Jan Södersten, Dividend Taxation, Share Repurchases and the Equity Trap, May 2009

2653 Assaf Razin and Edith Sand, Migration-Regime Liberalization and Social Security: Political-Economy Effect, May 2009

2654 Yin-Wong Cheung and Hiro Ito, A Cross-Country Empirical Analysis of International Reserves, May 2009

2655 Bart Cockx and Bruno Van der Linden, Flexicurity in Belgium. A Proposal Based on Economic Principles, May 2009

2656 Michael Melvin, Lukas Menkhoff and Maik Schmeling, Exchange Rate Management in Emerging Markets: Intervention via an Electronic Limit Order Book, May 2009

2657 Susanne Neckermann, Reto Cueni and Bruno S. Frey, What is an Award Worth? An Econometric Assessment of the Impact of Awards on Employee Performance, May 2009

2658 Steven Brakman, Harry Garretsen and Charles van Marrewijk, Economic Geography within and between European Nations: The Role of Market Potential and Density across Space and Time, May 2009

2659 Giovanni Facchini and Cecilia Testa, Reforming Legislatures: Is one House better than two?, May 2009

2660 Carsten Kowalczyk and Raymond Riezman, Trade Agreements, May 2009

2661 Oliver Falck, Stephan Heblich and Elke Luedemann, Identity and Entrepreneurship, May 2009

2662 Christian Lessmann and Gunther Markwardt, One Size Fits All? Decentralization, Corruption, and the Monitoring of Bureaucrats, May 2009

2663 Felix Bierbrauer, On the Legitimacy of Coercion for the Financing of Public Goods, May 2009

2664 Alessandro Cigno, Agency in Family Policy: A Survey, May 2009

2665 Claudia M. Buch and Christian Pierdzioch, Low Skill but High Volatility?, May 2009

2666 Hendrik Jürges, Kerstin Schneider, Martin Senkbeil and Claus H. Carstensen, Assessment Drives Learning: The Effect of Central Exit Exams on Curricular Knowledge and Mathematical Literacy, June 2009