**PROGRAMME**

**Wednesday 27 August**

|  |  |
| --- | --- |
| **TIME** | **ACTIVITY**  |
| 10.00-14.00 | Registration *Building 67* |
| 12.30-13.30 | Lunch *Garden Court Restaurant End* |
| 13.30-13.45 | Opening*Room 67/1027* |
| 13.45-14.45 | Keynote speaker 1: Roel Bosker (University of Groningen):Effectively dealing with differences between students - what do we know? *Room 67/1027* |
| 15.00-16.30 | Parallel session 1 |
| *Room 67/1003*Paper session 1: Improving reading outcomes | *Room 67/1007* Paper session 2: Student monitoring, data literacy and data use | *Room 67/E1001* Roundtable session 1: Alternative perspectives on educational effectiveness and improvement |
| 16.30-17.00 | Coffee break *Garden Court Garden End* |
| 17.00-18.30 | Parallel session 2 |
| *Room 67/1003*Paper session 3: Teacher effectiveness: quality and equity | *Room 67/1007* Paper session 4: Value added and educational effectiveness: international perspectives. | *Room 67/E1001*Paper session 5: Teachers’ work: impact of collaboration, data-based decision-making and performance management. |
| 19.00 | Reception and finger buffet *Garden Court Garden End* |

**Thursday 28 August**

|  |  |
| --- | --- |
| **TIME** | **ACTIVITY** |
| 9.30-10.00 | EARLI SIG 18 business meeting *Room 67/1027* |
| 10.00-10.30 | Coffee break *Garden Court Garden End* |
| 10.30-12.00 | Parallel session 3 |
| *Room 67/1003*Paper session 6: Effectiveness & equity: international perspectives | *Room 67/1007*Paper session 7: Teacher Development and professional learning  | *Room 67/E1001*Paper session 8:Theoretical, critical and alternative perspectives on educational effectiveness |
| 12.00-13.00 | Lunch *Garden Court Restaurant End* |
| 13.00-14.00 | Keynote speaker 2: David Reynolds (University of Southampton): The Future of Educational Effectiveness and Educational Improvement- Findings from the New International Handbook of Educational  Effectiveness and Improvement*Room 67/1027* |
| 14.00-15.30 | Parallel session 4 |
| *Room 67/1003*Paper session 9: Methodological advanced in educational effectiveness research | *Room 67/1007*Paper session 10: Leadership, school and system improvement | *Room 67/E1001*Paper session 11: Effects of homework, ICT and truancy in secondary and higher education |
| 15.30-16.00 | Coffee break *Garden Court Garden End* |
| 16.00-17.30 | Parallel session 5 |
| *Room 67/1003*Paper session 12:New developments in teacher effectivenessresearch | *Room 67/1007*Paper session 13:System effectiveness: international comparisons | *Room 67/E1001*Paper session 14: Teacher effectiveness, Subject knowledge and value added |
| 17.45-18.45 | Keynote speaker 3: Sam Stringfield (University of Cincinnatti):Toward Reliable, Sustainable School Reform: How far we’ve come and a glimpse at the road ahead.*Room 67/1027* |
| From 19.30 | Conference dinner (optional)*Cenos Restaurant, Highfield, Southampton* |

**Friday 29 August**

|  |  |
| --- | --- |
| **TIME** | **ACTIVITY** |
| 8.30-9.30 | Keynote speaker 4: Katharina Maag-Merki (University of Zurich):Toward Reliable, Sustainable School Reform: How far we’ve come and a glimpse at the road ahead*Room 67/1027* |
| 9.30-10.00 | Coffee break *Garden Court Garden End* |
| 10.00-11.30 | Paper session 6 |
| *Room 67/1003*Symposium 1: Marrying rigour and relevance: towards effective education for kindergarten teachers | *Room 67/1007*Paper session 15:Learning environments: equity and effectiveness | 7 *Room 67/E1001*Paper session 16: Improving schools and learning: Research and practise |
| 11.30-12.30 | Lunch *Garden Court Restaurant End* |
| 12.30-13.30 | Keynote speaker 5: Pam Sammons (University of Oxford):Combining rigour and relevance: the potential of mixed methods in educational effectiveness research*Room 67/1027* |
| 13.30-13.45 | Closing and departure *Room 67/1027* |

**PARALLEL SESSIONS: CONTENT AND ABSTRACTS**

**Wednesday, 27 August, 15.00-16.30. Parallel session 1**

**Paper session 1: Improving Reading Outcomes. Room 67/1003**

*Paper 1: Pascal Bressoux and Jérémy Pouille: Influence of school anxiety on reading comprehension performance.*

Abstract: This study aims at modeling the impact of school anxiety on students’ reading comprehension, integrating the mediating role of self-efficacy beliefs. 527 French students in 28 classes have been followed up from the beginning until the end of their grade-5 school year. Multilevel models show that students are more or less prone to anxiety depending on the class they attend. Multilevel models also reveal a nonlinear relationship between school anxiety and reading comprehension performance. The relationship is positive for low levels of school anxiety (performance gradually increasing for students who feel anxiety symptoms occasionally rather than never) but it is negative for high levels of school anxiety (performance gradually decreasing for students who feel often or systematically anxiety symptoms). This relationship is partially mediated by self-efficacy beliefs in reading comprehension.

*Paper 2: Mechteld van Kuijk, Marjolein Deunk, Roel Bosker and Lieneke Ritzema: The effect of a teacher professional development program on students’ reading achievement.*

Abstract: Currently, there are concerns on the reading comprehension results of Dutch students in primary school. In this study, we investigated whether student reading comprehension could be improved with help of a teacher Professional Development (PD) program targeting goals, data use, and effective instruction. All three components have shown to be positively related to student performance. As the long-term effects of early acquired reading skills have been documented in the literature, our PD program specifically targeted teachers of the second and third grades (student age: approximately 7 to 9 years old). In total, 33 teachers participated in the program which was conducted in 2011-2012 (40 hours in total). A pretest posttest control group design was used to examine the effect of the PD program on national standardized second- and third-grade reading comprehension assessments. The propensity score matching-approach was applied in order to match the classes in the experimental condition to equivalent control classes from a larger pool of “possible controls” – attained via a larger research conglomerate. Multiple classroom characteristics were taken into account during this matching procedure. After assessing the quality of the matching results (following recommendations in the literature), we conducted multilevel regression modeling to identify the effect of the program. Students in the experimental condition (n = 420) scored significantly higher on the standardized assessment than students in the control condition (n = 399), with an effect size of d = .37. No differential effects of the PD program were found in relation to initial reading performance or grade. We checked for the robustness of our results using different model specifications, and found similar albeit smaller results indicating that students in the experimental group were more than half a year ahead of the control group students.

*Paper 3: Jeannette Kunst, Thoni Houtveen and Wim van de Grift: How teachers can strengthen their students' reading performance by influencing their reading motivation.*

Abstract: This study among 21 teachers and their 547 eight years old students focusses on the effects of teaching behaviour on reading outcomes through the influence of motivation. Students’ reading level was explained by retention and task oriented behaviour and their time on task, and by teaching behaviour consisting of classroom management, fostering self-confidence of students, making sure that students read 30 minutes a day self-chosen age adequate books, book introduction and modelling literate behaviour, coaching the student’s reading process, reviewing with students what they have read, and the following elements of standards based teaching: monitoring of type of books read, setting reading goals, reflection on given instruction and construction and implementing intervention plans for students lagging behind. Further analyses reveal that teachers can enhance their student’s task oriented behaviour with efficient classroom management, making sure that their students read 30 minutes a day self-chosen age adequate books and reviewing with their students what they have read.

*Paper 4: Thoni Houtveen, Saskia Brokamp and Wim van de Grift: What teacher behaviour matters in improving reading performance?*

Abstract: About 25% of the pupils leaving Dutch primary education do not reach sufficient reading proficiency. Against this background a field experiment was conducted in 25 primary schools to improve reading performance in primary education. During three years, two times a year, 66 teachers teaching fluency to 9 year old pupils (grade 3) were observed during a silent reading lesson and coached in teaching skills. Every year teachers had to fill out a questionnaire about their teaching and response to intervention. During three years, the teachers received feedback and coaching on their development in these skills. After 3 years teachers’ growth in observed behaviour varied between a half and two standard deviations. Teacher growth in self-reported behaviour on response to intervention varied from almost zero to more than a standard deviation. At the start and at the end of the experiment, 1737 pupils took a standardized reading aloud test to measure their reading levels. At the start of the experiment 67,8% of the pupils reached sufficient reading proficiency. At the end of the experiment 86.0% of those pupils reached sufficient reading proficiency. Further analysis shows that pupils reading results at the end of the experiment are explained by pre measurement reading level, intelligence and age, and by the following teacher behaviour: classroom management, fostering self-confidence of pupils, book introduction and  modelling of literate behaviour, coaching the reading process, monitoring of type of books read, setting reading goals, reflection on given instruction, implementation of the core curriculum, implementation of the supplemental program for students lagging behind, constructing and implementing intervention plans.

**Paper session 2: Student monitoring, data literacy and data use. Room 67/1007**

*Paper 5: Marieke van Geel, Trynke Keuning, Jean-Paul Fox and Adrie Visscher: Data literacy development during a DBDM intervention.*

Abstract: Data literacy is assumed to be one of the preconditions for effective implementation of data based decision making (DBDM) in schools. The Focus-project is a two year training course aimed at implementing DBDM in primary schools in the Netherlands. Participants (teachers and school leaders) acquire knowledge and skills related to data use throughout the intervention. In the Focus-project, participants take a data literacy pre-test and post-test. Three different software programs are commonly used for analyzing student achievement data in the Netherlands, therefore three versions of this data literacy test were developed, using a set of general items for all participants, and software-specific questions for each of the systems. By means of multivariate IRT-analysis, data literacy development during the intervention can be modeled across student monitoring systems. Results are not yet available since data collection will be finalized in July 2014.

*Paper 6: Emmelien Akky van der Scheer and Adrie J. Visscher: The effects of an intensive one-year data-based decision making intervention for grade 4 teachers on students’ mathematical achievement.*

Abstract: In response to supposed declining student achievement levels the Dutch government has introduced its data-based decision making (DBDM) policy which is meant to promote teachers’ use of student monitoring system data for evaluating student progress, and the provision of tailor-made instruction to all students. Despite the growing demand from policy makers towards schools to work in data-based ways, empirical evidence of its effects is still scarce. The current study is a randomized controlled trial of the effects of a DBDM-intervention on student achievement. The research question is: What is the effect of an intensive DBDM-training for grade 4 teachers on students’ mathematical achievement? Grade 4 teachers of 60 Dutch primary schools participated in this project. The schools were randomly assigned to an experimental group (41 teachers) and a control group (36 teachers). The teachers in the experimental group received an intensive DBDM training, consisting of seven meetings and four coaching moments.Student performance was measured by means of standardized tests from the Dutch CITO student monitoring system. Since students are nested within classes, a multi-level model will be estimated to analyze the results. Post-test data will be available in June 2014, after which the analysis will be carried out.

*Paper 7: Marjan Faber: The Effects of Digital Student Monitoring Systems on Student Achievement.*

Abstract: Digital student monitoring systems (DSMS) are used in many schools in the United States and in European countries. Governments have high expectations; using DSMS is expected to result in higher student achievement. A meta-analysis was conducted to examine whether there is a significant effect of DSMS-use on student achievement. Fifteen experimental studies with 40 different effect sizes were included in the analyses. Effect sizes were computed using Cohen’s d or Hedges’ g. Comprehensive Meta-analyses was used for computing study weights, the average effect size, and for computing the effects of potentially influencing factors. The findings indicated DSMS-use has a large (effect size = 0.4) effect on achievement if teachers use the system for 2 till 4 students. Using DSMSs at the school-, or the school board-level has a small impact (effect size = 0.06) on student achievement. Factors that play a role in accomplishing effects are a high feedback frequency, feedback that includes instruction advice, and DSMSs that were combined with interventions in which teachers learned how to translate the feedback into tailor-made instruction. Analyses also indicated that the largest effect sizes were found in special education, and in studies with a short duration. No significant results were found for using feedback reflecting the results of the whole group of students, or in combination with benchmarks. On the contrary, not using benchmarks resulted in higher achievement. These findings support that feedback from DSMSs improves student achievement when teachers use the system for 2 till 4 students. Further research is needed to determine how, and whether the positive effects for small students groups can also be accomplished with larger groups of students.

**Roundtable session 1. Alternative perspectives on educational effectiveness and improvement. Room 67/E1001**

*Paper 9: Jane P. Preston: Education for Aboriginal Peoples in Canada: A Holistic Overview of Practice and Theory*

Abstract: One of the latest reports to surface from the federally-commissioned Canadian Panel on First Nation Elementary and Secondary Education for Students on Reserve provided a clear message: the quality and provision of education for Canada’s Indigenous people requires vast improvement (Haldane, Lafond, & Krause, 2011).  Additional research verifies that there is great disparity between in the quality and type of education that Aboriginal and non-Aboriginal students receive.  Considering this information, a blatant theme becomes clearly apparent: Aboriginal students need to be supplied with better educational resources and opportunities within the kindergarten to postsecondary education systems. Having articulated the need for improvement, there are palpable indications that positive changes are beginning to unfold with Aboriginal education in Canada.  The purpose of this article is to highlight the programs, services, and leadership issues that are having a positive impact on Aboriginal education in Canada and analytically reflect these milestones through a holistic model of a four-quadrant circle, with the concept of ethical space at its core.  More specifically, the Aboriginal programs presented within four thematic areas relate to: (a) early childhood education; (b) Aboriginal pedagogy, language, and culture; (c) governance and partnerships; and (d) postsecondary education.  Within the round-table discussion, specific programs, initiatives, and policies are named and described in each of the four areas. After these practical examples are explicated, a holistic view of continued improvement to Aboriginal education is analyzed via the model of the four-quadrant Medicine Wheel and Ermine’s (1997. 2000, 2007) concept of ethical space.

*Poster 1: Beatrice Riser-Lembang: Emotions and Learning in Early Childhood*

Emotions and the emotional dimension in learning is getting more and more attention and in the focus of research and teaching. Science of Education, Social Science, Psychology and Neuroscience have demonstrated in recent years that there is a close connection and/or interaction between emotions and cognitions. The subject matter is complex, as it affects the human being as a whole; that is, his body, his soul, and his embeddedness in a social context as well as the survival strategies (approaches); and last but not least, the strive for a happy, healthy and well balanced life. The emotional relationship is a prerequisite (must) for learning.

In the poster, the real exchange between emotions and learning and the effect of emotions on the learning process will be represented and discussed.

**Wednesday 27 August, 17.00-18.30. Parallel session 2.**

**Paper session 3: Teacher effectiveness: quality and equity. Room 67/1003**

*Paper 10: Charalambos Y Charalambous and Ermis Kyriakides: Capitalizing on Generic and Content-Specific Teaching Practices to Explain Student Performance: Insights from TIMSS and PISA Secondary Analyses.*

Abstract: Recent years have seen concerted efforts to more rigorously examine instructional quality, with a plethora of frameworks and/or classroom observation systems been advanced toward this end. A closer examination of these tools suggests that scholars have largely considered either generic or content-specific teaching practices. Capitalizing on one generic and one content-specific tool, this study makes a case for attending to both types of practices. Analyzing data from TIMSS and PISA studies, we examined the added value of considering both generic and content-specific practices in explaining variance in student performance in mathematics. The analyses for the entire sample showed that more variance was explained when considering both types of practices rather than when focusing on either type; important insights were also gleaned when examining selected individual countries. Implications for the design of international comparative studies are drawn.

*Paper 11: Zhenzhen Miao and David Reynolds: A comparative study on the effectiveness of mathematics teaching in England and China.*

Abstract: This paper focuses on the quantitative part of a cross-national study on the effectiveness of mathematics teaching to the age of 10. Stratified sampling method was applied in two countries, England and China. The stratum of participants was defined by two criteria at the school level: 1) pupils with medium level mathematics performance, and 2) pupils from medium level socioeconomic background. As a result, 10 teachers from England and 12 teachers from China with their pupils were recruited for this study. The main purpose of the quantitative part of the study is to evaluate teaching and learning and correlate observed teacher behaviours with pupil learning outcomes – mathematics performance and perceptions of mathematics teaching and learning. 22 mathematics lessons in each of the classrooms were observed and video-recorded with triple cameras focusing on the whole class, the teacher and pupil groups; pupils (n=560) from 19 classes completed a 40-item test paper derived from TIMSS 2003 maths test at two points in time; pupils (n= 691) from all 22 classes answered a 74-item questionnaire measuring alternative learning outcomes – perceived teaching and learning. Lessons were measured with two internationally validated observational systems. Two tests were marked against the marking guidance of TIMSS 2003, and mean scores were calculated at the classroom and then the country levels, at which were average scores of each domain of the pupil questionnaire calculated as well.  To correlate teacher behaviours and two types of learning outcomes, Pearson’s r correlation coefficients were calculated, and F-test conducted. Results indicate that on average Chinese pupils outscored their English peers by 28 points out of 100 in test 1 and 21 in test 2, and Chinese teachers scored higher on most effective teaching scales than their English colleagues. Details can be found in the report.

*Paper 12: Evi Charalambous, Leonidas Kyriakides and Bert P.M. Creemers: Promoting Quality and Equity in Education: The Impact of School Learning Environment.*

Abstract: This paper investigates the impact of School Learning Environment (SLE) on student achievement gains (quality) and on reduction of variation in the initial student outcomes (equity). Student achievement in mathematics and language of all grade 5 students (n=2503) of 50 primary schools were measured at the beginning and at the end of school year 2004-2005. A teacher questionnaire was used to measure the SLE. A follow-up study took place in the same 50 schools during the school year 2008-2009. Qualitative characteristics of policy for improving the SLE were found to be associated with both quality and equity. Implications for research and practice in promoting quality and equity in education are drawn.

**Paper session 4: Value added and educational effectiveness: international perspectives. Room 67/1007**

*Paper 13: Sally Thomas and Wen Jung Peng: Time Trends in Schools Value Added Performance and the Influence of Teachers Professional Development on Student Outcomes in China.*

Abstract: This paper presents selected findings from two UK DFID/ESRC funded studies: Improving Educational Evaluation and Quality in China (IEEQC) and Improving Teacher Development and Educational Quality in China (ITDEQC). The paper seeks to provide new insights about (i) the impact of student, teacher and school characteristics and other factors on students attainment and progress at senior secondary school, (ii) the relevance of these factors in the evaluation of school effectiveness in China and (iii) patterns in school effects across 4 student cohorts (2009-2012) and 3 regions (east and west). The methodology involves exploring the nature and extent of school effectiveness in three separate Chinese regions. Innovative quantitative methodology (multilevel modelling) is used to analyse longitudinal datasets comprising students’ examination, prior attainment and background information as well as teacher/school process and context data from 130+ schools, 17,000+ teachers and 300,000+ students. The findings replicate, validate and extend previous similar research (using data from only one student cohort) that indicated significant differences in value added measures of school effectiveness between schools and regions in China. The paper also investigates the existence and relevance of professional learning communities in Chinese schools and identifies links between teacher development and students’ “value added” progress and outcomes. The findings are discussed in terms of educational policy and practice in mainland China and internationally, focusing on teacher development and school improvement in different contexts, and the implications for school evaluation and accountability.

*Paper 14: Bernardita Munoz Chereau: The hierarchical structure of the educational system: a multilevel analysis of Educational Effectiveness in Chilean secondary education.*

Abstract: This paper explores the multilevel nature of pupil achievement data in the Chilean school system. The analysis of a longitudinal sample of a cohort of 177,461 students, nested within 7,146 classrooms within 2,283 schools within 43 LEAs, that took the National System of Measurement of Educational Quality tests (SIMCE) in two subject outcomes (Spanish and maths), in grades 8 and 10 in 2004 and 2006, and their family questionnaires, is analysed using 2, 3 and 4 level Multilevel Model (MLM).  In order to explore the extent to which pupil intake and background features of Chilean secondary schools influence students’ performance, 4 types of approaches recommended to study school effectiveness (OECD, 2008) are also employed: Raw, Contextualised Attainment (CA), Value Added (VA) and Contextualised Value Added (CVA) models. The results will be presented and compared with studies carried out in developed countries (England and China), in an effort to situate the findings in the broader knowledge-base of Educational Effectiveness Research (EER). The findings replicate, validate and extend previous research exploring the hierarchical nature of the educational system in Chile (students within schools) that showed significant differences in value added measures of school effectiveness between secondary schools. Hence, the novel features of this article are the exploration of different dimensions of EER (LEA, school, class and student) in a new context (Chile); the use of better methodological approaches underused in Chile (MLM and VA/CVA), and the comparison of the results with relevant studies conducted in developed countries. The findings are discussed in terms of the importance of considering the hierarchical structure of the educational system in Chile, the implications for the new school evaluation and accountability policy recently introduced in the country and the validity of EER in Latin-America in general.

*Paper 15: Tomasz Zoltak: Value-Added Indicators of School Effectiveness: The Limits of Interpretation.*

Abstract: In the late 1990s, when value-added indicators of school effectiveness were introduced in England and the United States, they were thought to give a precise answer to a specific question: what is the contribution of a given school to its students' achievements? With time, some doubts regarding the validity of value-added indicators started arising but many researchers considered them as temporary difficulties that can be easily overcome. However, an in-depth analysis leads to the conclusion that we should be very cautious when making decisions based on the value-added indicators. This paper shows the limitations of possible interpretation of the value-added indicators when used as: 1) a tool to assist teacher assignments to students, 2) part of the evaluation system, 3) support for parents when choosing a school and 4) a tool for self-evaluation for teachers. I carefully analyze the limitations resulting from: 1) the properties of the statistical methods used to estimate the value-added indicators (with a focus on the independence between fixed effects and random terms), 2) the model specification (omission or inclusion of control variables) and 3) the empirical properties of the indicators (low stability from year to year). The analyses are based on the different implementations of the value-added systems of American TVAAS/EVAAS, English Value-Added and Contextual Value-Added indicators, and Polish Educational Value-Added indicators.

**Paper session 5: Teachers’ work: the impact of collaboration, data-based decision-making and performance management. Room 67/E1001**

*Paper16: Anett Wolgast, Natalie Fischer and Désirée Theis: You are not alone – Cooperation and Collegiality as Resources to Decrease Teachers’ Workload.*

Abstract: Experienced workload of teachers can affect the quality of classroom instruction and in turn students’ motivation (e.g. Klusmann et al., 2006). However, it is supposed that teachers experience less workload if they feel supported by their colleagues (Smith & Gillespie, 2007). In line with this assumption the model of instructional quality of teachers (Klusmann et al., 2008, p. 703) indicates that colleagues’ support serves as a resource for teachers and has a positive influence on their “instructional performance”. Based on the contact hypothesis (Aronson et al., 2014) can be assumed that frequent cooperation of teachers may lead to more colleagues’ support. However, teacher cooperation is generally rather low (Gallimore & Goldenberg, 2001). To our knowledge, up to now, the relationship between teacher cooperation, collegiality and workload is explored rarely. This paper aims to close this gap by answering the following questions: a) How is experienced workload related to colleagues’ support and high cooperation frequency? b) Assuming that high colleagues’ support could be explained by frequent cooperation, a mediation hypothesis is tested: Is experienced workload explained by frequent cooperation via increased colleagues’ support? Analyses are based on data from the longitudinal “study on the development of all-day schools” (StEG) in Germany. The sample consists of *N* = 6,311 teachers. Dependent variable is experienced workload. Without including colleagues’ support in the analysis the findings show a significant correlation between experienced workload and cooperation frequency. Including colleagues’ support, there is no significant correlation between experienced workload and cooperation frequency. However, experienced workload correlates significant with colleagues’ support and this correlates significant with cooperation frequency. Thus, there is a mediation effect. Therefore, payed cooperation times could increase cooperation and collegial support between teachers and decrease their experienced workload.

*Paper 17: Trynke Keuning, Marieke Van Geel and Adrie Visscher: The impact of a data-based decision making intervention on schools’ social networks.*

Abstract: While the importance of teacher collaboration has been foregrounded in many studies on DBDM, our knowledge of what this collaboration looks like and how it may be influenced by a DBDM-professional development intervention is limited. The current study takes a social network perspective to explore how collaboration takes shape in the interactions among teachers as they engage in a DBDM-reform. Three social networks characteristics (density, reciprocity and centralization) of 32 Dutch elementary schools regarding three DBDM-topics are examined at two time points as they engage in a two-year DBDM-professional development reform: discussing student achievement, discussing the achievement goals that teachers would like to accomplish, and discussing instructional strategies to accomplish the goals set. Multivariate multiple regression analysis was used to analyze the differences between the network structures at the start of the intervention on the one hand, and after having participated in a DBDM-professional development intervention one year on the other hand. Findings show that at the start of the intervention DBDM-collaboration social networks are relatively small (especially for discussing performance goals), and centralized around a few influential persons within the school (especially for discussing student achievement). Results also show that social networks of weakly collaborative school teams at the start become stronger during the DBDM-intervention, which is reflected by increased network density and increased network reciprocity.

*Paper 18: Kim Bellens, Jan Van Damme, Wim Van Den Noortgate and Sarah Gielen: Effects of performance management of teachers on educational quality and equity.*

Abstract: Two main goals of education are quality and equity, i.e. establish high achievement and assure that all students are to the same extent able to attain their highest possible achievement level. Many educational systems introduced past few decades more practices of performance management of teachers to reach both aims. However, to investigate the effects of these practices on educational quality and equity, research mostly makes use of correlational methods, which implies that research on the causal influence of different practices of PM on quality and equity is scarce. To make more profound causal conclusions on the effect of different practices of performance management of teachers, this study uses a difference-in-difference design. We investigate the relationship between changes in teacher performance management and changes in quality and equity over time (between 2003 and 2011) in two different groups (Grade 4 and Grade 8) over 46 countries. Data of TIMSS are used to answer the research question. In the analyses, aggregation of school level information to the country level reduces selection bias to a minimum. Next to that, omitted variables are largely excluded by looking at evolutions over time. Quality is operationalized in this study as countries’ mean achievement score on math respectively science. Equity is indicated by the relationship between student social and ethnic background and student’s achievement score on math respectively science. The effect of four different practices of performance management of teachers is investigated, i.e. evaluating teachers by means of (1) observation by the principal or senior staff, (2) observation by inspectors or other persons external to the school, (3) student achievement and (4) teacher peer review. The results of this study will be presented at the conference. With the statistical methods used in this study, we are able to give more causal indications of the effect of several practices of performance management of teachers on educational quality and equity.

**Thursday 28 August, 10-30-12.00. Parallel session 3.**

**Paper session 6: Effectiveness & equity: international perspectives. Room 67/1003**

*Paper 19: Lorena Ortega, Lars-Erik Malmberg and Pam Sammons: School effects on Chilean children’s achievement growth in language and mathematics: an accelerated growth curve model.*

Abstract: The present study investigates school effects on student achievement growth in Chile. In order to do so the shape and predictors of primary students’ achievement trajectories in language (Spanish) and mathematics are examined and the magnitude of school effects is estimated using a contextual value-added approach. The study’s data sets were obtained by linking data from Chilean assessment programmes and administrative records and feature an accelerated longitudinal design comprising participants in 4 overlapping cohorts, together spanning Grades 3 to 8 (N = 10,534 and 11,403 students for language and mathematics, respectively). Results indicate non-linear upward growth on student achievement in primary school and significant individual differences in both, achievement status and rate of development over time. Boys were found to perform better in mathematics and worse in language. In mathematics, this gender gap increases as students progress through primary school. Low-SES students have growth trajectories that differ from that of other students although, in language, they seem to somewhat make up for some of their initial arrears. Other variables that predicted student achievement status were number of books at home and age. School effects on students’ growth trajectories were found to be sizeable and larger than those found in previous studies using similar model specifications in industrialised economies. At the school level, variables that predicted achievement status were school achievement mean, school achievement standard deviation in the relevant subject. School sector was not a significant predictor either of student achievement status or of achievement growth, after controlling for the student and school characteristics previously mentioned. These finding align well with the school effectiveness literature as school differences with respect to the achievement growth of their students are bigger than usually demonstrated in gain scores models and larger school effects are found for student achievement growth than for status.

*Paper 20: Lei Zhang: Teacher and School Effectiveness of Senior High Schools in Western China.*

Abstract: This paper uses secondary data with a three-level hierarchical structure (Student within teachers within schools) from the Improving Teacher Development and Educational Quality in China (ITDEQC) project to conduct school and teacher effectiveness research for senior high schools in one Local Education Authority (LEA) in western China.  The findings showed that there are marked differences between schools and teachers in their effectiveness. Schools and teachers are not egalitarian to all initial achievers. The most and least “advantaged” students seem to make more progress than the average students, which suggests that the average students appear to be ignored by schools and teachers. The effects of prior attainment on student’s achievement also differ for students with different citizenship status or number of siblings.

*Paper 21: Pam Sammons, Kati Toth, Kathy Sylva, Edward Melhuish, Iram Siraj, Brenda Taggart and Rebecca Smee: Influences on students’ academic attainment and progress in secondary school in England*

Abstract: For 17 years, the Effective Pre-school, Primary and Secondary Education Project (EPPSE) has contributed to the international debate on what influences student outcomes. EPPSE as a longitudinal study followed 3000 children from pre-school until secondary school and beyond. This paper focuses on the relationships between a range of individual student, family, home, pre-, primary and secondary school characteristics and students’ academic attainment in year 11 at secondary school (age 16). It also compares the latest findings with those found for students’ attainment at younger ages, highlights the influences of secondary school on students’ attainment in the core curriculum areas and studies their academic progress across secondary school between the ages of 11 and 16. We used multiple measures for academic outcomes[1], both continuous (total GCSE score, English and maths GCSE grades) and dichotomous (achieved 5 or more exams at grades A\*-C including exams in English and maths). Analyses using MLM revealed that differences in attainment related to background influences which emerged early (at age 3) have remained fairly stable through to the end of year 11. Both mothers’ and fathers’ (to a lesser extent) educational levels strongly predicted academic attainment. Female students obtained higher total GCSE scores, higher grades in English and were more likely to pass 5 or more exams including English and maths with grades higher than C when compared to male students. Other significant predictors of academic attainment included family salary, parents’ highest SES, the early years home learning environment (HLE) and free school meal (FSM) status. Additionally, the effects of pre-school, primary and secondary school experiences continued to predict students’ academic attainment in year 11, even after controlling for background characteristics.

*Paper 22: Henry Nsubuga Kiwanuka, Jan Van Damme, Dickson Nkafu Anumendem and Speranza Namusisi: Multilevel analysis of the factors affecting mathematics achievement of First Year Secondary school students in Uganda.*

Abstract: This study explores the sources of variability in mathematics achievement for Ugandan students at the student-, classroom- and school-level, based on the international research literature. The mathematics scores and questionnaire responses of 4819 First Year secondary students(about 14-15 years) from 78 classes of 49 schools were analyzed. A three-level multilevel linear modeling was used. The results indicate that out of the total variance in mathematics achievement, the relevant factors could explain 68.49%, 14.07% and 17.44% , respectively, of  student-, classroom- and school-level differences. For student-level factors, gender, prior mathematics achievement, attitude toward mathematics were significant predictors of achievement. For classroom-level factors, class-mean of prior math achievement and of students' perception of good classroom assessment were significant predictors. For school-level factors, school-mean of students' perception of good classroom assessment and of prior math achievement had significant association with math achievement.

**Paper session 7: Teacher development and professional learning. Room 67/1007**

*Paper 23: Marina Pinskaya, Irina Grunicheva and Sergey Kosaretsky: Professional Learning Community (PLC) in Challenging Schools as a Tool to Improve Teaching and Learning. First Russian Experience.*

Abstract: This study contains the description of the very first outcomes of the application of the model of Professional Learning Communities (P.Senge, M.Fullan, A.Harris, M.Jones) in Russian Schools. But these first outcomes are worthy for presentation for several reasons. Firstly because the study describes one of the first (if not the first) experience of introducing this model into Russian school and thus can be considered as an innovation. Secondly because this innovation was implemented not in advantaged high-performing schools (which is typical of post-Soviet Russian reform traditions) but in disadvantaged low-performing schools working in challenging social and economical context. Thirdly because in Russia there are many successful practices when high quality of education is achieved by the attraction of external recourses, but practices of mobilization and effective use of internal school resources are in great demand. And this study shows that PLC can become effective tool of building school internal capacity and ensuring sustainable improvement. The main research questions in relation to PLC were as follows: can PLC model be used in context of Russian school practice; can PLC model serve as a basis of improvement programs for low-performing schools in challenging circumstances; and do domestic traditions of school culture and domestic strategies of school management allow to introduce PLC model as efficiently as it has been introduced in other countries. During the research the following methods were used: semi-structured interviews with teachers and administration; questioning of teachers; questioning of pupils; classroom observations.

*Paper 24: Trynke Keuning, Marieke Van Geel, Adrie Visscher and Jean-Paul Fox: Teacher Quality development during an Intervention aimed at Data-Based Decision Making.*

Abstract: An intervention to implement and sustain data-based decision making (DBDM) in primary schools was executed in 42 primary schools in the Netherlands. The successful implementation of DBDM requires a high level of professionalism of teachers. In order to test whether teacher quality changed and improved during the DBDM-reform, teaching quality was measured by means of student perceptions in grades two to six. The quality of approximately 400 individual teachers was rated four times over the course of two years (on average 7,500 students per measurement occasion; on average 20 students per teacher, per measurement moment). By means of multivariate multilevel growth analysis, changes in teacher quality will be modeled. The results of 3 measurement occasions are available, the last measurement will take place in June 2014. Exploratory analysis of collected data shows that teachers improved in the extent to which they ‘challenge students’, and teach in ‘goal-oriented’ ways.

*Paper 25: Jantine Kuijpers, Thoni Houtveen and Wim van de Grift: Jantine Kuijpers, Thoni Houtveen, Wim van de Grift.*

Abstract: Teachers’ professional development is an important issue in effectiveness research as teachers are the key figures when it comes to influencing student performance. What teaching skills should be implemented in schools to reach results on student level becomes more and more clear from the literature (e.g. Hattie, 2012). In the school improvement and school effectiveness literature the role of the external facilitator in teachers’ professional development is presupposed. Remarkably, little research has been done on the effects of external facilitator behaviour on supporting teachers’ professional development. To fill in this gap, in a previous study, we developed a model for effective training of the desired teaching skills and attitudes: the integrated professional development model (IPD-model). The model is developed on the basis of school improvement literature on the one hand and teachers’ professional development literature on the other hand. In the present study is examined to what extent external facilitators carry out the diverse (sub)categories of the IPD-model and what (sub)categories of the model explain the degree of implementation of an innovation and teachers’ attitude and concerns with respect to an  innovation. The results show that the external facilitators on average, carry out at least 60% of the activities, belonging to the distinguished (sub)categories of the IPD-model. The outcomes of multilevel regression-analyses show that facilitators influence more strongly teachers’ concerns with respect to the innovation (23% explained variance) than that they influence teachers’ attitude with respect to the innovation and degree of implementation (8% and 10% explained variance). Some of the (sub)categories from the IPD-model explain the greater part of variance on the implementation of the innovation and teachers’ attitude and concerns with regard to the innovation. Among these (sub)categories are ‘creating a sense of urgency’ and ‘monitoring and evaluating the school improvement project with the teaching team’.

**Paper session 8: Theoretical, critical and alternative perspectives on educational effectiveness. Room 67/E1001**

*Paper 26: Jaap Scheerens and Sigrid Blömeke: Operational models of educational effectiveness, what can be learned from path analyses.*

Conceptual models of educational effectiveness tend to assume that higher level educational conditions support or stimulate lower level conditions in influencing educational outcomes. Path analytic modeling makes direct and indirect "causation" visible and can also accommodate the multi level structure of the data. The operational path models can be seen as instruments that link conceptual models with empirical data. This presentation is based on review studies and meta-analyses conducted by the first author, and recent studies in the field of teacher education effectiveness conducted by the second author. Summary results based on these studies are used to answer the following questions:

* What are the major independent variables used in these studies?
* Which intermediary school and teaching variables appear to be successful in manifesting indirect effects?
* What is the relative importance of direct and indirect effects in the examples that were studied?
* To what extend do the operational models and empirical results support the structure and choice of variables in the available conceptual models?
* How can the results of this review be related to more general theoretical perspectives on educational effectiveness and ineffectiveness?
* Do the results have anything to say on greater parsimony of models and the relative success of more general strategies to enhance effectiveness and school improvement?
* What is the particular contribution of models that have tried to open the black box of teacher education effectiveness by studying connections between opportunities to learn provided in teacher education, teacher knowledge and effective teaching strategies?

*Paper 27: Lawrence Angus: School Effectiveness and Educational Reform: A Sociological Critique.*

Abstract: In 1993, the author published a review essay on ‘the sociology of school effectiveness’ in the *British Journal of Sociology of Education*(Angus 1993). The review, which was widely cited in ensuing years, argued that the school effectiveness movement was in need of serious sociological critique. Since then, the school effectiveness field has made substantial advances in relation to its analytical methods, diversity of perspectives and approaches, and stronger recognition of the importance of social context. Despite these advances, however, I argue in the paper that serious criticisms of school effectiveness remain. Moreover, the easy absorption of school effectiveness discourse and priorities into the prevailing neoliberal educational policy regime has had the effect of bestowing legitimacy on educationally damaging policies that embrace and promote undue competition, marketization, performativity and managerialism. The paper concludes with a discussion of alternative ways of ‘doing’ educational reform, particularly in relation to issues of equity and social justice.

*Paper 28: Megan Walberg: Advocacy for Mental Health Indicators in Models of Academic Achievement*

Abstract: There are many indicators used to account for academic achievement, however, some areas, such as mental health are deserving of greater attention than currently given. The aim of this paper is to illustrate the importance of pupils’ mental health (in particular, emotional and behavioural difficulties) in predicating achievement by using data from the ROOTS British longitudinal study (N=1038). Specifically, a series of regression models were conducted to select key variables from major domains (the family, school, and individual) used in the literature for predicting academic achievement. Variables with the largest effect sizes were entered into a single regression model to determine whether mental health remained a significant predictor of academic performance. Mental health (β = -2.71, p <. 001) and the additional indicators included in the model (prior achievement in English, β = 34.31, p <. 001; prior achievement in math, β = 26.95, p <. 001; mother’s education, β = 15.97, p <. 001; teacher quality, β = 1.64, p <. 001; and special educational needs, β = -39.10, p <. 001) significantly predicted academic achievement, with the exception of one variable (adverse family environment, β = -3.74, p = .27). A series of regression models were conducted using a subsample (n= 248) of ROOTS participants, who had completed a set of cognitive tasks to explore the potential mechanisms linking mental health factors and achievement outcomes. The results suggested cognitive impairments were present for those with high levels of mental health problems, and that there were possible sex-differentiated pathways: males with mental health problems appeared to have challenges in risk-evaluation tasks, where as females with mental health difficulties struggled with emotionally valenced, decision-making tasks. The overall findings will be discussed in the context of their implications for policy-makers and school practices.

**Thursday 28 August, 14.00-15.30. Parallel session 4**

**Paper session 9: Methodological advances in Educational Effectiveness Research. Room 67/1003**

*Paper 29: Ioulia Televantou: Phantom Effects in School Composition and Self-Concept Research: Consequences of Failure to Control Biases Due to Measurement Error in Traditional Multilevel Models.*

Abstract: The present paper, a methodological-substantive synergy, consists of two distinct, but inter-related studies. Study 1 seeks to quantify the impact of failing to account for measurement error on school compositional effects of prior mathematics achievement as used in value added models to explain relative school effects. Multilevel structural equation models are incorporated to control for measurement error and/or sampling error. Analysis involving a large sample of English primary students in years one and four, reveals a significant and negative compositional effect on students’ subsequent mathematics achievement that becomes more negative (twice as large in magnitude) after controlling for measurement error. Further analyses with the same set of data (Study 2), demonstrates a negative compositional effect of school average mathematics achievement on subsequent mathematics self-concept, verifying the Big Fish Little Pond Effect (BFLPE) for students as young as five to seven year old. Adjustments for measurement and sampling error result in more negative BFLPEs. Mediation analysis reveals that the prevalence of BFLPEs explains some part of the negative compositional effects of achievement detected in Study 1. My findings have important methodological, substantive, and theoretical implications for on-going debates on the school compositional effects on students’ outcomes, because nearly all previous research has been based on traditional approaches to multilevel models, which are positively biased, due to the failure to control for measurement error.

*Paper 30: Christian Bokhove: Demonstrating the consequences of not taking into account sampling designs with TIMSS 2011 data.*

Abstract: The topic of comparative international large-scale assessments (LSA) has always had a lot of attention from policy makers and educational researchers, inviting criticism. One criticism concerns the fact that the complex sampling design of LSA is not always taken into account. This paper aims todemonstrate the consequences of not taking into account the sampling design of one such assessment, TIMSS 2011.Three features, weights, proficiency estimation with plausible values and variance estimation with jackknife are used in single level (students) and multilevel (students and schools) cases. The results show that the consequences can be significant, but are not completely in line with previous literature*.*

*Paper 31: Daniel Caro: Causal mediation in educational research: An applied example using international assessment data.*

Abstract: The paper demonstrates an application of causal mediation in educational research. The applied example draws on data from the Progress in International Reading Literacy Study (PIRLS) 2011 to examine the mediating role of early literacy activities in the relationship between family socioeconomic status (SES) and reading attainment. PIRLS has a cross-sectional design, which in theory would prevent us from studying mediation and causal questions. The paper argues, however, that because the questions on early literacy activities refer to the period before the child entered school and family SES is relatively time-invariant, the research design with retrospective data satisfies mediation conditions. That is, it is argued that the mediator (i.e., early literacy activities) precedes the outcome (i.e., reading attainment) and that the treatment (i.e., family SES) precedes the mediator and the outcome. Further, the rich set of variables available in PIRLS allows us to control for treatment and mediator confounders. The causal mediation approach integrates traditional mediation analysis with the potential outcomes framework. Sensitivity analysis establishes benchmarks for interpreting mediation results causally. Preliminary results indicate small mediation effects that are sensitive to violations of causality assumptions. The value of international assessment datasets for studying causal effects and causal mechanisms is discussed.

*Paper 32: Keith Zvoch and Joseph J. Stevens: The Identification of Program Effects using Piecewise Growth Models in a Comparative Interrupted Time Series Design.*

Abstract: Piecewise growth models (PGM) were used to estimate and model changes in the pre-literacy skill development of kindergartners in a moderately-sized school district in the U.S. Pacific Northwest (N = 356). PGM were applied to interrupted time series (ITS) data that arose within the context of a Response-to-Intervention (RtI) instructional framework. During the kindergarten year, upwards of 16 literacy assessments were conducted and 30 minutes of daily supplemental instruction was initiated with struggling readers to promote the attainment of literacy benchmark goals. The use of PGM provided analytic flexibility to specify a discontinuous learning function and test a number of within and between group model parameters related to the evaluation of program efficacy. Results revealed that the rate of literacy acquisition statistically increased following the initiation of the instructional supplement during the 7th week of the academic year, but absolute and relative gains in literacy performance were more modest and variable during the second half of the academic year. More generally, the demonstration shows that PGM techniques provide a flexible and powerful approach for mapping the structure of RtI instructional practice while simultaneously documenting the immediate and more distal responses to intervention. The specificity with which the initiation, withdrawal, and subsequent re-initiation of supplemental instruction at key points in the academic calendar (including the onset and end of winter break) can be modeled offers researchers and practitioners a means to more clearly identify normative development as well as patterns of inadequate response. Key issues in applying PGM including evaluation of functional form, choosing "clock" codes that correctly represent time functions and hypotheses of researcher interest, and how to properly interpret PGM model results are addressed.

**Paper session 10: Leadership, school and system improvement. Room 67/1003**

*Paper 33: Alma Harris, Michelle Jones, Donnie Adams and Jacqueline Perera: Explaining Educational Effectiveness: The Art of Implementation Science.*

Abstract: **This paper looks at high performing education systems in Asia through the lens of leadership and leadership development. It argues that the best educational systems scientifically build the leadership capacity for improvement as part of an implementation science geared to maximizing performance. Drawing upon initial scoping work from a cross-national study, the article focuses upon two high performing systems, Singapore and Hong Kong. The paper concludes by proposing that the top performing education systems in Asia not only create the leadership capacity to consistently outperform others but also invest in an implementation science that defines, delineates and ultimately, determines exceptional performance.**

***Paper 34: James Ko:*** *Arbitrariness in theory building: the glory and price of parsimoniousness.*

Abstract: This paper explores the arbitrariness in theory building and argue that the golden principle of parsimoniousness in scientific research has its cost. Based on the data of a study examining the dynamic impacts of school principals on changes in student learning through school improvements, a series of structural educational models are set up based on a set of variables related to principal values, principal efficacy, policy enactment, and leadership practices addressing rival theories such as transformational leadership and instructional leadership. Evidence for the study was provided by 168 senior teachers of 41 schools to a survey, along with value-added student achievement data between 2006 and 2008. Path analytic techniques were adopted to address the objectives for the study. Results indicated that parsimoniousness should not be the only criterion for evaluating rival models. Models set up to test the relative significance of constructs showed the shortcomings of leadership theories that failed to address policy enactment. Constructs under-researched like values and qualities of principals were found to be an important antecedent of principal efficacy. School leaders’ efficacy also had strong, positive impacts on the enactment of education reforms and school improvements.

*Paper 35: Mariane C. Koslinski, Fatima Alves, Gregory Elacqua and Humberto Santos: Accountability pressure and teachers practices: evidences from Rio de Janeiro.*

Abstract: A large body of research has documented the effects of pressures coming from school accountability systems. On one hand, some studies identify impacts in terms of improvement in student achievement. On the other hand, part of the literature argues that this apparent improvement is due to principal and teacher ‘gaming’ the system. Nevertheless, there has been little attention paid to the meaningful educational changes in school policies and teacher practices resulting from accountability pressures.  The lack of evidence in mainly due to a dearth of survey data on principal and teacher behavior in response to accountability. This paper analyzes the impact of a recent accountability program implemented in Rio de Janeiro –  Prêmio Anual de Desempenho[Annual Performance Bonus]– and provides evidence on the effects of accountability systems on teacher practices and school results. We analyze a survey conducted of a census of public schools in Rio de Janeiro, coupled with administrative data on student achievement and instructional policies and practices.  We also conduct a more detailed survey of a smaller sample of 200 schools in Rio de Janeiro to better understand principal and teacher decisions in response to accountability pressures.  We show that low performing schools that face accountability pressures changed their instructional practices in meaningful ways.   We do not find significant effects on teaching practices, suggesting that many of the changes are implemented top-down by principals, without involving teachers in the process.  In addition, we find evidence of the effects of accountability pressures on student achievement, which we maintain are likely to be related to the changes in instructional policies we identify in the school surveys.

**Paper session 11: Effects of homework, ICT and truancy in secondary and higher education. Room 67/E1001**

*Paper 36: Kati Toth, Pam Sammons, Kathy Sylva, Edward Melhuish, Iram Siraj and Brenda Taggart: Homework and academic attainment and progress in English secondary schools.*

Abstract: The relationships between the time students report to spend on homework, their self-regulation and academic attainment and progress in English secondary schools (year 11, at age 16) are explored. Although the links between homework and academic outcomes have been the subject of research in many countries and different phases of education, the conclusions are not always convergent. A meta-analysis of the U.S. research showed a modest positive effect of homework on academic outcomes, stronger in the middle and high school (Cooper, Robinson, & Patall, 2006). Recently, more complex statistical models have explored the relationship between homework and academic progress in maths (Trautwein, Koller, Schmitz, & Baumert, 2002; Trautwein, 2007). These studies found that at school level there was a positive relationship between homework and progress in mathematics, but at student level this relationship was negative. The present paper used MLM and SEM to investigate the relationships between homework and academic achievement and progress using GCSE[1] results from English secondary schools. Multilevel analyses showed that self-reported time spent on homework was a statistically significant, positive and moderately strong predictor of academic attainment and progress even when students’ individual (age, gender, birth weight etc.), family (SES, FSM, salary) and home learning characteristics were controlled. Additionally, moderately strong effects for time spent on homework were obtained even when the students’ perceptions of their school’s emphasis on learning was also controlled. SEM modelling explored possible causal relationships between time spent on homework and academic outcomes in year 11 and the role of prior self-regulation (from year 6). The models identify important direct and indirect relationships between gender, mother’s qualification level, self-regulation, time spent on homework and academic outcomes.

*Paper 37: Jane P. Preston, Brittany A. E. Jakubiec, Rachel Earl, Julie Jones, Angela Poirier, Kaitlyn MacKenzie, Joseph N. Lillo, Shannon Kemp and Robyn Christensen: GoogleDrive in Higher Education: Experiencing Transformation Learning.*

Abstract: The purpose of this article is to describe undergraduate student and instructor experiences when incorporating GoogleDrive into a Bachelor of Education (BEd) course.  Participants for this qualitative study were 8 first-year Bachelor of Education (BEd) students and their instructor.  Participants provided written answers to 16 open-ended questions and participated in two focus group interviews.  Findings showed that by using GoogleDrive, participants experienced a transformation in attitude and skill with regard to these technological platforms.  Analyzed through the concept of transformational learning theory, the constructivist learning experienced by students and the instructor enhanced collaboration and communication between one another.  An implementation of the study is that in order for BEd teacher candidates to incorporate technology into future Kindergarten to Grade 12 classrooms, they need to gain confidence and skills with using technology during undergraduate courses.

*Paper 38: Wei Zhang and Chang Zhu: Students’ Perceptions of Blended Learning Comparing With Traditional Learning in Higher Education.*

Abstract: Blended learning sounds like a nice idea--mixing a traditional classroom environment with online components--but is it actually effective? (David Nagel, 2008). In order to examine the effectiveness of blended learning comparing to the traditional face-to-face learning mode, and also explore the possible reasons, this research investigated the students’ perceptions of blended learning comparing with traditional learning through a questionnaire which is designed based on the Seven Principles For Good Practice in Undergraduate Education.(Arthur W. Chickering and Zelda F. Gamson, 1987). The seven principles are Activity, Expectation, Cooperation, Interaction, Time, Feedback, and Diversity. The questionnaire was distributed to students in Beijing Union University and 652 questionnaires were collected, including 399 blended learning students and 253 traditional learning students. SPSS was used as the tool to analyze the data. Independent samples T-test was mainly used to compare the means of scores students have answered for the questionnaire. The results showed that students got higher perception scores for blended learning than traditional learning for all the seven principles, in which Activity, Expectation, Cooperation, Interaction, Time, and Diversity had significant differences between blended learning and traditional learning, while Feedback did not have significant difference between blended learning and traditional learning. The results do not only prove that blended learning are more effective than traditional learning, but prove that the reasons why blended learning are more effective than traditional learning are that blended learning stimulate students’ activity and expectation, promote cooperation among students and interaction between students and teachers, and provide diversity and time efficiency for students. Future research may be conducted in other higher education institutions in China or in other countries.

*Paper 39: Christine Saelzer: Missing mathematics classes: Does truancy affect students' mathematical literacy?*

Abstract: Large-Scale-Assessments such as PISA aim on measuring basic student competences, also known as literacy. The relationship of student competence and their performance in tests is commonly expressed in a probabilistic way, e. g. by means of the Rasch-Model (Adams & Wu, 2007). A crucial, but yet hardly examined question in this context is *where* students actually acquire their competences – At school? Or rather somewhere else? This problem is especially relevant for students who at least occasionally miss classes and therefore have relatively little exposure to guided learning at school. This study takes into account that lessons take place only for students who are present and asks about the relationship between class attendance and competence. Skipping classes or skipping school is generally referred to as truancy or truant behavior (Author, 2010). The association of class attendance in mathematics and students’ mathematical literacy is therefore guiding this paper. Research questions are: To what extent is students‘ mathematical literacy associated with individual, class-level and school-biographical characteristics?; and Is students‘ mathematical literacy related to the frequency of missed classes in mathematics? In order to examine the research questions, stepwise hierarchical linear regression models have been specified. Mathematical literacy was predicted using individual background- and school-biographic characteristics as well as truancy and class-level predictors. The data used for this study stem from a national grade-based sample of 400 grade-9-classrooms participating in PISA 2012 in Germany. Furthermore, the PISA student questionnaire has been enhanced by nationally added questions concerning students‘ individual school attendance behavior. N = 9998 grade-9-students completed the questionnaire. Results showed that mathematical literacy is not independent from individual student characteristics and students‘ personal school-related biography. On the other hand, and going beyond individual characteristics, mathematical literacy in PISA is related to individual class attendance and the frequency of absences in a student’s class (mean absence rate). These aspects as well as the variability of the role of the students’ grade in mathematics within this predictive pattern suggest that the debate on competence and performance should consider a behavioral level as well—not only from an individual student’s point of view, but also with regard to the class a student is sharing lessons with. Possible implications for school and teaching will be discussed using the PISA 2012 data.

**Thursday 28 August, 16.00-17.30. Parallel session 5**

**Paper session 12: New developments in teacher effectiveness research. Room 67/1003**

*Paper 40: Pam Sammons, Alison Kington, Ariel Lindorff-Vijayendran and Lorena Ortega. Inspiring Teaching: What we can learn from exemplary practitioners.*

Abstract: This paper discusses findings from a small-scale, mixed methods study of ‘inspiring’ teaching. The study, commissioned and funded by CfBT, included case studies of a purposive sample of 17 primary and secondary teachers in England who were nominated by their head teachers as exemplary practitioners whose practice could be viewed as inspirational for their colleagues and pupils. The entire project was comprised of two phases. In the first phase, experienced inspectors visited schools to observe and interview teachers and speak with their pupils. The second phase, on which this paper focuses, was conducted by a research team from two major UK universities. The mixed methods design for this phase of the project included integration of qualitative and quantitative approaches across the data collection, analysis, and interpretation stages. Quantitative data collection methods included two research-based structured classroom observation schedules, specifically the International System of Teacher Observation and Feedback (ISTOF) and the Lesson Observation Form for Evaluating the Quality of Teaching (QoT), ranking grids listing a variety of teacher characteristics, and pupil surveys. Qualitative data sources included semi-structured teacher interviews, qualitative notes from classroom observations, and one open-response pupil survey item. Several common themes emerged across the teacher perspectives, pupil responses, and researcher observations, indicating that inspiring practice may relate to the following: positive relationships, good classroom management, positive and supportive classroom climate, formative feedback, enjoyment, and a high quality learning experience overall. Additionally, evidence from this study suggests that inspiring and effective practice are complementary; effective practices may facilitate inspiring teaching, though inspiring practice is linked to additional features beyond those associated with effective teaching. These findings have both theoretical and practical significance. Besides informing understandings about what ‘inspiring’ teaching means, the results of the project will also be disseminated among CfBT schools and teachers.

*Paper 41: Jasmin Decristan, Eckhard Klieme, Mareike Kunter, Jan Hochweber, Gerhard Büttner, Benjamin Fauth, Lena Hondrich, Svenja Rieser, Silke Hertel and Ilonca Hardy: Interaction Between Embedded Formative Assessment and Classroom Process Quality in Promoting Student Learning.*

Abstract: Principles of effective teaching refer to generic factors of teaching, specific teaching practices, and teaching principles of the content matter. This study contributes to research on educational effectiveness by examining the interplay between a specific teaching practice (i.e., curriculum-embedded formative assessment) and general characteristics of classroom process quality (i.e., student ratings of cognitive activation, supportive climate, and classroom management) in promoting student learning. To contribute to this research, we used data from an intervention study conducted in primary school science classes. In this paper, we compare two of four intervention conditions, i.e. embedded formative assessment (17 classes) and control group (11 classes). To control for teaching of the content matter, all of the 28 teachers used a scripted science unit. Findings add to research on effective teaching and embedded formative assessment, which like each dimension of classroom process quality proved to affect students’ science understanding. Most importantly, classroom process quality positively moderated the effectiveness of embedded formative assessment.

*Paper 42: Alexander Naumann, Svenja Rieser, Gerhard Büttner and Eckhard Klieme: Effects of cognitive activation on the probability of solving cognitively demanding items.*

Abstract: In 2009, Klieme, Pauli and Reusser proposed a three-dimensional model of teaching quality, distinguishing supportive climate, classroom management, and cognitive activation. Numerous studies provided empirical evidence that cognitive activation is positively related to students’ test scores (e.g., Lipowsky et al., 2009). However, to date it is relatively unclear which responses to specific items are affected by the degree of cognitive activation in class. In our study, we expect that experiencing a high level of cognitive activation in class should not only relate to higher test scores, but also to solving cognitively more demanding items. Using explanatory multilevel IRT models, we found empirical evidence that cognitive activation not only is related to solving more items, but also impacts the probability of solving cognitively more demanding tasks. Our results underline the role of cognitive activation as an important aspect of teaching quality, fostering students’ conceptual understanding.

**Paper session 13: System effectiveness: international comparisons. Room 67/1007**

*Paper 43: Cesare Aloisi and Peter Tymms: The stability of PISA scores and the impact of education policies on them.*

Abstract: This paper seeks to identify the extent to which, and over what time frame, national education policies could be said to affect PISA scores. Firstly, the implications of the high correlations observed between PISA scores from 2000 to 2012, and the trends in the correlations, are considered. Secondly, the role of education policies for the emergence of such trends is discussed. Thirdly, multilevel models are fit to the observed scores. These include: growth models with time as a predictor variable, autoregressive models, as well as growth models with country-level explanatory variables as additional covariates. Results are analysed and evaluated: the findings suggest thatany effect of policy making should be generally considered as (a) restricted to few jurisdictions and (b) gradual and progressive. Therefore, the authors advise that cautions should be exercised when suggesting that the impact of national education policies can be seen in less than a decade - approximately the span of four administrations of PISA.

*Paper 44: Jenny Lenkeit and Daniel H. Caro: Education System Effectiveness: Linking Educational Effectiveness Research and International Large-Scale Assessments.*

Abstract: Reports of international large-scale assessments tend to evaluate and compare education system performance based on absolute scores. And policymakers refer to high performing and economically prosperous education systems to enhance their own systemic features. But socioeconomic differences between systems compromise the plausibility of those comparisons and references. The paper applies conceptual and methodological approaches from Educational Effectiveness Research to investigate how effectively education systems perform and how effectively they change their performance over time by accounting for socioeconomic differences between systems and cohorts (assessment cycles). Data from four cycles of the Programme for International Student Assessment (PISA) is analysed. Results indicate that the quality of systems is evaluated differently if assessed by absolute performance scores or effectiveness measures. The study contributes to methodological developments of effectiveness research in international large-scale assessments and provides relevant information for policymakers to further look into policies, structures, and reform measures that have favoured effectiveness.

*Paper 45: Maria Georgiou and Leonidas Kyriakides: Investigating the impact of national educational policy on student achievement: a longitudinal study in six European countries.*

Abstract: This paper investigates the extent to which national educational policy on teaching and school learning environment can explain differences between six European countries (Cyprus, Belgium/Flanders, Germany, Greece, Ireland, and Slovenia) and schools within countries in promoting student learning outcomes. In each country, a sample of at least 50 schools was drawn and tests in mathematics and science were administered to all Grade 4 students (n=10742) at the beginning and at the end of the school year 2011-2012. For the construction of the tests, permission was obtained from IEA to use the released items of TIMSS 2007. Evidence on the national policy for teaching and school learning environment of each country were collected through content analysis of the official policy documents.  The national/state educational policy that originates in the official policy documents of the nation/state in the form of educational law and guidelines, may present differences from the one that is promoted by the different actors. To investigate the policy in the way that is promoted by the stakeholders, semi-structured interviews were also conducted with individuals occupying key positions in the different educational systems. Multilevel analysis revealed that most system level factors of the dynamic model explain student achievement gains in mathematics and science. Implications for research and design of the international studies are drawn.

**Paper session 14: Teacher effectiveness, subject knowledge and value-added. Room 67/E1001**

*Paper 46: Aleksandra Jasinska and Anna Hawrot: Does an effective school teach intelligence? Validity of Polish value-added indicators.*

Abstract: Assessment of school effectiveness requires good measurement tools.  One such a tool - value-added indicators (or educational value-added - EVA) – is being developed in Poland. The aim of the study was to assess the validity of EVA indicators for Polish lower secondary schools (LSS). We assumed that school effectiveness correlates with numerous processes and phenomena related to students’ social, emotional and cognitive development. Thus, we expected EVA estimates to correlate with pupils’ relative intellectual growth during LSS. We analyzed data of over 5200 students  from 291 classes and 150 schools gathered when pupils were in grade 1 and 3 of LSS to test the hypothesis. The analyses indicated the existence of a correlation between EVA indicators and relative mean intelligence gains among students.

*Paper 47: Benjamin Caspar Fauth, Jasmin Decristan, Svenja Rieser, Eckhard Klieme and Gerhard Büttner: Teacher popularity as a predictor of student achievement.*

Abstract: The contribution examines effects of teacher popularity on student learning. Teacher popularity refers to the affectively tinged general impression of the teacher as perceived by his or her students. Previous research indicates that teacher popularity can be a useful indicator for teacher effectiveness, especially in primary school. Our empirical analyses focus on student learning in a standardised teaching unit using a longitudinal pre-post measurement design (N = 1070 students, 54 classes). Effects of student rated teacher popularity were estimated in multilevel regression analyses controlling for individual intake variables and student rated teaching quality. Results show a significant effect of teacher popularity at the classroom level albeit not the individual level. Teacher popularity is thus crucial for student learning, not as a characteristic of the individual student-teacher relationship but as an attribute of teachers as judged by their students. We discuss the relationship between teacher popularity and teaching quality with regard to theoretical models of educational effectiveness.

*Paper 48: Henk van den Hurk: Does teacher knowledge matter in improving pupils’ reading fluency?*

Abstract: We studied to what extent teachers’ pedagogical content knowledge and teachers’ instructional behaviour are accountable for the variance in pupils’ reading achievement. This study involved 120 teachers, teaching reading fluency to 9 to 12 year old pupils (grades 3 to 6) in 17 primary schools. The variables on the teacher level included: ‘pedagogical content knowledge of reading’, ‘modelling literate behaviour’, ‘coaching and supporting the reading process’ and ‘fostering feelings of competence and self-confidence in pupils’. Multilevel regression analyses with two levels were performed, with different measures of pupils’ reading achievement as the dependent variable. The results of this study, ultimately available by august 2014, can be of importance for the future curriculum development in teacher training institutions and in courses for staff development.

**Friday 29 August, 10.00-11.30. Parallel session 6.**

**Symposium 1: Marrying rigour and relevance: towards effective education for kindergarten teachers. Room 67/1003**

Chair: Prof. Dr. Sigrid Blömeke, Humboldt-University of Berlin, Germany

This symposium examines prospective kindergarten teachers’ competence. Based on Weinert (2001), the term “competence” is defined as a domain-specific multidimensional latent trait that includes cognitive abilities and affective-motivational facets underlying performance in real-world situations. During the past 10 years, substantive research on prospective primary and secondary teachers’ competence has been carried out (e.g., TEDS-M; Blömeke et al., 2011, 13; Tatto et al., 2012).

However, a huge research gap exists concerning the structure, the level and the development of kindergarten teachers’ competence during teacher education, especially in the field of mathematics (National Advisory Panel, 2008) although we know that the development of children’s mathematics ability depends on pedagogical support during kindergarten (Van Oers, 2009). The study “KomMa – Competence of Kindergarten Teachers in Mathematics” funded by the German Federal Ministry of Education intends against this background to develop, firstly, a concept model of kindergarten teachers‘ competence that includes the cognitive abilities and affective-motivational dispositions necessary to master real-world situations.

*Paper 49: Blömeke, S., Baack, W., Tengler, M. and Wedekind, H: Effects of opportunities to learn on prospective kindergarten teachers’ mathematics pedagogical content knowledge*

Abstract: Children’s mathematical ability develops already during their early years. This development strongly depends on the quality of the pedagogical support provided, among others by kindergarten teachers (van Oers, 2009). To what extent opportunities to learn (OTL) mathematics or the teaching of mathematics are part of kindergarten teachers’ training and what level of knowledge the teachers have reached at the end of their training has not yet been examined sufficiently (National Advisory Panel, 2008). The study presented here examines these OTL of prospective kindergarten teachers and their effects on the teachers’ mathematics pedagogical content knowledge (MPCK).

*Paper 50: S. Dunekacke, L. Jenßen, M. Grassmann & S. Blömeke: Prognostic validity of kindergarten teachers’ mathematics pedagogical content knowledge.*

Abstract: Kindergarten teachers have to diagnose and foster learning processes of children in an informal environment (van Oers, 2009). There is a lack of studies focusing on the structure, level and development of kindergarten teachers’ competence during teacher education, especially in the field of mathematics (National Advisory Panel, 2008). The present study consists of a sample of n=354 prospective kindergarten teachers. To examine how disposition and performance work together two instruments are used. First, an achievement test for MPCK, developed within the KomMa project containing 36 items. Second a video-based assessment with 24 predominantly open-ended items to record “perception of situations” and “planning on actions”. The videos show daily life-situations from a day-care center and were selected based on an expert-rating. Data collection took place on two days within a period of two weeks. The research questions were examined by structural equation modeling, two models were compared. The result can be regarded as a first indicator of the prognostic validity of the achievement test developed in KomMa and of the relevance of MPCK for performance in Kindergarten. Since we used a latent-variable framework, the effects can be evaluated as moderate support of the prognostic validity of the achievement test (Kane, 2013).

*Paper 51: Jenßen, L., Dunekacke, S., W. Baack & Blömeke, S.: The Relationship of Kindergarten Teachers’ Mathematics Knowledge and Mathematics Anxiety: An Application of Latent State-Trait Theory.*

Abstract: Educational concepts of competence assume that competence is a trait. For example, Weinert (2001) defines competence as the ability to successfully master problems in variable situations. Evidence suggests that mathematical competence is such a stable personality characteristic (Aunola et al., 2004). However, Mischel (1968) assumed that occasion-specific influences may also affect the measurement of constructs. Latent state-trait theory (LST) allows the investigation of situation-specific effects on obtained achievement scores and to separate them from dispositional factors (Steyer, Schmitt & Eid, 1999, Geiser & Lockhart, 2012). Math anxiety has been discussed as a predictor of varying relevance for effectiveness in the field of mathematics (Ma, 1999). Research revealed that math anxiety consists of trait amounts and occasion-specific parts, if systematic manipulation of the situation exists (Goetz et al., 2013).The purpose of this study is to close the research gap with respect to kindergarten teachers by applying LST. Furthermore, it is intended to clarify the relationship of math anxiety and mathematical competence in terms of LST. The mathematical competence of 350 future kindergarten teachers from Berlin and Lower Saxony were tested twice within three weeks with a mathematics test developed in the project KomMa. A series of structural equation models was applied to examine the research questions. Results indicate that there are no significant occasion-specifics in mathematical competence and math anxiety. Correlations between these two constructs on trait level were of moderate negative size.

**Paper session 15: Learning environments: equity and effectiveness. Room 67/1007**

*Paper 52: Spyros Konstantopoulos: Class Size Effects on Mathematics Achievement in Cyprus: Evidence from TIMSS 2003 and 2007.*

Abstract: The effects of class size on student achievement have gained considerable attention in education research and policy the last 30 years (e.g., Blatchford, 2003; Finn & Achilles, 1990; Nye, Hedges, & Konstantopoulos, 2000). The present study examines the effects of class size on mathematics achievement using 8th grade data of the 2003 and 2007 samples of the Trends in International Mathematics and Science Study in Cyprus. Class size effects on student achievement in Cyprus have not been documented in the literature and thus our study aspires to fill in that gap. Our hypothesis is the class size reduction will increase student achievement. We employed two-level models to capture the clustering of students within schools. The main outcome was mathematics achievement and the main predictor was class size. Other student (e.g., gender, socioeconomic status or SES), teacher (e.g., education, experience), and school variables (e.g., size, composition) were included in the models as covariates. We also used instrumental variables (IV) to facilitate causal inferences about class size (Angrist & Lavy, 1999). To construct the instrument we used a nation-wide rule about maximum class size in Cyprus which since 2000 has been 25 for middle-school and high-school grades. The results indicated overall that the association between class size and mathematics achievement was not different than zero in 2003 and 2007 both in two-level and IV models. That is, we could not detect evidence of class size effects on mathematics achievement. Thus in Cyprus class size may not be a school resource that can improve student achievement at the middle school level. We observed gender differences in mathematics however that favored female students both in 2003 and 2007. Family SES as expected was positively and significantly related to mathematics achievement both in 2003 and 2007. All other variables were not related to mathematics achievement.

*Paper 53: Katja Scharenberg, Wolfram Rollett and Wilfried Bos: The importance of differential learning environments for the development of achievement in reading, mathematics and reasoning.*

Abstract: In most countries, some kind of tracking is applied to handle students’ diverse abilities. Tracking aims at a homogenisation of learning groups which is assumed to facilitate teaching and learning. In Germany, tracking already starts from the beginning of secondary school when students are at the age of 10 or 12. Previous studies showed that secondary tracks offer different opportunities for the students’ learning development often disadvantaging those in the lower tracks. On the one hand, these disadvantages are caused by track-specific study programmes and leaving certificates. On the other hand, also the composition of the different learning environments with regard to students’ social background and abilities leads to a track-specific development of achievement. Some studies showed effects of differential learning environments on the development of domain-specific achievement such as reading and mathematics. There is, however, only little evidence of tracking effects on the development of the students’ non-verbal reasoning ability. The aim of our contribution is a comparative analysis of the importance of compositional and tracking effects for the development in three different competence domains (reading, mathematics, non-verbal reasoning) using one dataset for a whole student population. We also analyse to what extent compositional and tracking effects are confounded. The analyses are based on data from the longitudinal study KESS in Hamburg (Germany). Multi-level models examine effects of student characteristics (N = 9,414) and aggregated variables on school level (N = 178).Multi-level analyses show that students in academic tracks have a higher development of achievement in all three domains. Moreover, schools with a higher average achievement composition offer a higher potential for the students’ development in all domains, whereas the social composition has an effect only on the development of reasoning. The decomposition of variance emphasises the common importance of both compositional and tracking effects.

*Paper 54: Michal Modzelewski and Jolanta Pisarek: Average students’ attentiveness and class size. Argos observation system in action.*

Abstract: The poster presents research results of the observational study that was conducted as a part of the Longitudinal School Effectiveness Study in Poland. We used the Argos observation system and its student-focused observation mode to collect our data. The observation study was conducted in 68 fifth-grade classrooms (children aged 11-12 years). Observers performed observations during eight Polish and seven Math lessons in each classroom, they observed eight students per lesson. We divided research sample into small (<= 19 students, 27 classes) and large classes (>= 23, 41 classes). Previous research shows some relationship between class size and the student engagement in class. To analyse this relation in reported study we used two specific behaviors: student works on task and student doesn’t work on task, which were registered by the observers during all observed samples (22-25 time samples per lesson). Results show students’ tendency to work more on task in small classes. During both Maths and Polish lessons there is significantly smaller average chance of observing student not working on a task in the small class than in the large one. Furthermore, students on average work more on task in small classes and this result was statistically significant for Maths lessons. Statistically significant dispersion of the average chances of observing student not working on task in large classes than in small ones was observed during Maths lessons as well.

**Paper session 16: Improving schools and learning: Research and practise. Room 67/E1001**

*Paper 55: Leonidas Kyriakides, Evi Charalambous, Athena Michaelidou and Bert Creemers: Promoting Student Learning Outcomes in Socially Disadvantaged Schools: The Impact of the Dynamic Approach to School Improvement.*

Abstract: This paper investigates the extent to which the Dynamic Approach to School Improvement (DASI) can promote student learning outcomes in schools with low socioeconomic status. A sample of 40 primary schools was selected. The schools were randomly allocated in two groups. Feedback on the functioning of the school factors was provided to all schools and schools of the first group (control group) were encouraged to develop their own action plans to promote student learning outcomes. Schools of the second group were asked to make use of DASI for improvement purposes. Student achievement in mathematics at the beginning and at the end of the school year was measured. Using multilevel modelling techniques, it was shown that schools which made use of DASI managed to promote student learning outcomes more than schools which formed the control group. No differential effect on students coming from different socioeconomic background was identified. Implications of findings are, finally, drawn.

*Paper 56: Marieke van Geel, Trynke Keuning, Jean-Paul Fox and Adrie Visscher: Assessing the effects of a school wide data-based decision making intervention on student achievement growth in Dutch primary schools.*

Abstract: Despite growing international interest in the use of data to improve education, relatively few studies examining the effects on student achievement are available. In the present study, the effects of a two-year data-based decision making intervention on student achievement growth were investigated. A total of 53 primary schools in the Netherlands participated in a project aimed at implementing data-based decision making throughout the entire school organization. Student achievement data was collected over the two school years prior to the intervention and during the two intervention years. Linear mixed models were used to analyze the differential effect of data-use on student achievement, controlling for background variables at the school and student level and accounting for individual growth in student achievement from grade three to eight. A positive mean intervention effect over students, schools and grades, and heterogeneity in school intervention effects was estimated. Heterogeneity in performance of students in the study prior to intervention and during intervention were not attributable to differences in observed student background variables. High intervention effects were identified for low-SES schools and students, leading to the conclusion that the data-based decision making intervention especially improved the achievement of students in low-SES schools.

*Paper 57: Elisabeth Dokalik-Jonak: Train the RIGHT brain: mnemonic device for ESL primary classes in Austria and Finland*

Abstract: Children of an ESL primary class acquire English with a remarkable speed, when taught by using mnemonic devices. Recent studies at the University of Teacher Education in Vienna have shown that six to ten year olds are capable of discerning 10 different sounds by just listening to "sound-jingles" and playing didactical games with "sound-cards". In three studies we investigate the effects of the above mnemonic devices on language production and accuracy. Two of these studies are completed and the data will be presented. However, the third and present study will also deal with the topic of long-term memory learning in an Austrian and Finish primary school class. The goal of this survey is to explore this topic, focusing on the data and arguments of the previous studies, and to contrast Austrian and Finish language learning in primary education.