

TÁMOP 3.1.9-11/1-2012-0001
„*Developing Diagnostic Assessments*” project

6TH SZEGED WORKSHOP ON EDUCATIONAL EVALUATION

ABSTRACTS

29-30 April 2014

Venue:

Szeged Committee of the Hungarian Academy of Sciences

7, Somogyi Street, Szeged

Institute of Education
Doctoral School of Education
Hungarian Academy of Sciences



ABSTRACTS

29th April 2014

Session A

<i>Mari-Pauliina Vainikainen</i>	Development of Learning to Learn Skills in Primary School
<p>In Finland, schools' effectiveness in fostering the development of transversal skills is evaluated through large-scale learning to learn (LTL) assessments. This article presents an overview of the development of these skills by showing with longitudinal data of 608 pupils how the cognitive component of LTL in fourth and sixth grade is related to students' cognitive competence and reading skills in the beginning of first grade. The development of learning-related motivational beliefs and their relation to the cognitive LTL competences are also followed. Finally, the cognitive competences and motivational beliefs are used in predicting pupils' performance in computer-based complex problem solving (CPS) tasks. The results show that the development of children's thinking and problem solving skills are not fully determined by earlier individual differences in cognitive competences. Motivational beliefs begin to play a role in explaining cognitive LTL performance gradually from age 10 on, and their effect is slightly stronger in the computer-based CPS tasks. It is concluded that the Finnish learning to learn model is a valid measure for predicting 12-years-olds' performance in a complex problem solving situation which is expected to be closer to the demands of their future work life.</p>	

<i>Sirkku Kupiainen</i>	Do Increasing Skills in Key Subjects Explain the Development of Students' Transversal Skills in Primary Education?
<p>In the Finnish framework, the transversal learning to learn competence is understood to comprise general cognitive competence and motivational attitudes which guide the use of that competence in the direction of the expectations of the school or of society at large, leading to lifelong learning (Hautamäki et al., 2002, 2010; see also Csapó, 2007). The development of both dimensions is understood to be built and fostered through good education in the curricular subjects. It has been shown that both the cognitive and the motivational component of learning to learn predict students' school achievement (Hautamäki & Kupiainen, 2014; see also Adey et al., 2007; Demetriou et al., 2011) but also that while students mid-term grade point average (GPA) is related to their attainment in the cognitive learning to learn tasks in the end of grade nine (Kupiainen et al., 2014), their motivational attitudes have no independent impact (or rather, a weak negative one) on their test attainment (unlike to the GPA), supporting the basic claim of the model. As the GPA already carries the impact of the motivational dimension of learning to learn in it, the claim of the Framework regarding the development of learning to learn through curricular learning can only be partially confirmed with such a model.</p> <p>To look at the role of curricular advancement in the development of learning to learn, the present study relies on not the GPA but on students' attainment in Finnish and Mathematics at two measurement points imbedded in a longitudinal study on the development of</p>	

learning to learn competence in Finnish basic education. In this, the present study presents an auxiliary view to the Vainikainen et al. (submitted / see Presentation 1) view on the development of students' transversal skills by focusing on the role of students' growing competence in Finnish and Mathematics in this development. To further elaborate the role of curricular learning in the fostering of learning to learn as one critical transversal skill of tomorrow, the study will also implement the same model to look at the development of complex problem solving as measure with the interactive computer-based MicroDYN tasks (Greiff et al., 2013).

<i>Risto Hotulainen</i>	Developmental Trajectories of Students' Mathematics Self-Concept and Cognitive Development in Secondary Education
<p>The ultimate goal of the educators is to have positive effect on learners' development and promote his or her realistic perceptions related to such development. Especially, changes of those responses which reflect perceptions of well-being, self-esteem and competences are of interest in the cognitive psychology because such perceptions are stated to reflect individuals' personal estimate about his or her attainment level. Many educational programs describe their objectives by enhancing positive and realistic self-perceptions and well-being. Especially study of such generalizable phenomena becomes interesting when knowing that, for example, real cognitive competences of individuals are increasing by their experiences, schooling and age but simultaneously their corresponding perceptions decrease. In this study the general trend line of the development of math self-perceptions and measured cognitive development of Finnish learners was studied. During a six-year longitudinal learning to learn study from grade 6 onwards, 589 Finnish students from the capital area were followed. Latent growth curve model along with the latent profile analyses were used to analyze effect of change based on three gathered time points. Results showed that the most of the self-perceptions changes of students which, indeed followed their cognitive competence development, were attributable by the changes of the educational settings referring to the frame of reference effect. Possible explanations of the findings, especially related to relation between math self-perception and cognitive competence are discussed.</p>	

Session B

<i>David Greger and Jana Straková</i>	Czech Longitudinal Study in Education (CLOSE) – Design and Preliminary Results –
<p>Czech Longitudinal Study in Education (CLOSE) is a 7-year research project (2012-2018) jointly pursued by the Economics Institute of the Academy of Sciences of the Czech Republic, the National Training Fund, and the Faculty of Education of Charles University in Prague. The project is financed by the Czech Science Foundation as a “Center of Excellence”. CLOSE undertakes a comprehensive research program in order to gain a better understanding of the key issues related to schooling, skill acquisition, and achievement. The team of the Faculty of Education focuses on the impact of tracking and school choice on the development of knowledge and skills of students in compulsory education. To study these issues it carries out two longitudinal surveys: the first one follows a representative sample of 1200 6 years old students at their transitions from pre-</p>	

primary to primary education; the second follows a representative sample of 3500 students participating in PIRLS and TIMSS 2011 at their transitions to lower and upper secondary education. In 2012, the data on students' aspirations and practices at the transition from primary to secondary education were collected from students at the end of grade 5. At the beginning of grade 6, students in different lower secondary tracks wrote tests in math, Czech grammar, reading and learning to learn, and filled out a questionnaire related to their motivations, school experience and home background. Next data collection will take place in grade 9 to estimate the value added of individual tracks. In 2013, a qualitative study on school choice was carried out among parents of six years old children. Recently a quantitative follow up is under preparation. The presenters will describe the methodology of the study and will briefly introduce the first results.

<i>Benő Csapó</i>	Theoretical and Methodological Frameworks of the Hungarian Educational Longitudinal Program
<p>The Hungarian Educational Longitudinal Program (HELP) was launched in 2003 to explore students' development in educational context. The aims of the program include describing developmental trajectories, modeling development, exploring the factors that influence development, studying the stability and fluctuation of development, identification of blocking factors causing failures and verifying the predictive validity of diagnostic instruments and several educational tests.</p> <p>Samples of the program are representatively drawn from the school population of Hungary (n>5000). The program started in 2003 with 3 cohorts: with 1st, 5th and 9th grade students. This design allowed covering the 12 years of education in three sections by collecting data for 4 years. Since 2007 in every four years a new 1st grade sample has been added to the program (so far in 2007 and 2011).</p> <p>The instruments of data-collection are cognitive and affective tests and questionnaires which are administered to entire classes at the beginning and at the end of the school years. A school readiness test battery is administered to the first graders entering the program in the first weeks of their first school year. The program focuses on three main domains: reading, mathematics and science. Beyond testing students achievements at these domains, several reasoning skills are assessed (e.g. inductive reasoning, problem solving), as well as some affective domains (e.g. motivation, future orientation, attitudes). In the first years, paper and pencil tests were used but some tests for the 2011 cohorts have been administered online. The next cohort joining in 2015 will be assessed by computerized tests delivered online. By using an encrypted student identification code, the data generated from HELP can be connected with the achievement data of the Hungarian National Assessment system in reading and mathematics measured at grades 6, 8 and 10.</p> <p>The results of HELP may be applied to a number of practical problems. Its experiences helped to establish the Development of Diagnostic Assessments project where students' achievement data will also be longitudinally connected.</p>	

<i>Gyöngyvér Molnár</i>	Developing Diagnostic Assessments: Review of Recent Results
<p>The project entitled „Developing Diagnostic Assessments” aims at developing the quality and effectiveness of the Hungarian educational system by providing regular feedback on students’ cognitive development in three main domains: reading, mathematics and science in the first six grades of primary school. The developmental work includes the focus shift of the frameworks for diagnostic assessments from paper-based to computer-based assessment in line with the electronically based item bank developed within the framework of the project. There were ca. 1000 computer based 4-5 item tasks for 1-to 6 graders in each of the main literacy domains developed in 2013. The tasks were allocated to 481 15-20 item clusters and the clusters were organised into test booklets composed of three clusters according to the specifically developed rotation design. The first large scale data collection (n_student=9267, n_measurement_point=12401) was carried out in December 2013 via the eDia platform. According to the scaling results the tasks were improved, reedited and further tasks were developed. The second large scale assessment with 478 booklets and over 25000 measurement points begun in spring 2014. Parallel with the item development on the main areas the eDia platform was constantly tested in our 500 partner schools using innovative tests of 14 cognitive and affective skills and competencies as well (e.g. ICT literacy, visual and musical skills, problem solving; n_measurement_point_2013>50.000). According to the requirements of the different competency fields, eDia was further developed. Now, with its state-of-art technology it is able to administer not only static, but dynamically changing items and is suitable not only for individual assessment, but for assessment in groups with immediate feedback as well.</p>	

Session C

<i>Yvonne Neeb, Johannes Naumann, Tobias Richter, Julia Knoepke and Maj-Britt Isberner</i>	Effects of Spoken Word Recognition at School Entry on Text Comprehension in Grade 3 are Mediated Through Lexical Quality
<p>Good text comprehension rests partly on the quality of lexical representations, i.e. reliable and tightly integrated representations of the phonological and orthographical properties of words. Little is known, however, about how the development of lexical quality is related to accurate spoken word recognition. Accurate spoken word recognition might help beginning readers to build lexical representations of words, thereby facilitating the acquisition of word reading and text comprehension skills. The aim of the present study was to investigate more precisely the development of spoken word recognition and its contribution to the development of reading comprehension.</p> <p>Following these considerations, the present research employed a longitudinal design to examine how the accuracy of spoken word recognition skills develops in beginning readers, and whether the effects of spoken word recognition on later text comprehension skills are mediated through the quality of phonological and orthographical representations of printed words. A sample of 346 German primary school children was tested at school entry for their spoken word recognition. At the end of Grade 1, Grade 2, and Grade 3, spoken word recognition and the quality of phonological and orthographical representations were assessed with computerized tests. In addition, at the end of Grade 3, text comprehension</p>	

was assessed with a standardized German reading test. Parallel tests were constructed for measuring the quality of phonological and orthographical representations and corresponding spoken word recognition skills. For all predictors, different but parallel item sets were used at different times of measurement.

Latent growth curve modelling indicated that spoken word recognition skill increased through the early years of primary school, following a quadratic trend. While spoken word recognition skill improved substantially between the beginning of Grade 1 and the end of Grade 2, only marginal gains were observed between the end of Grade 2 and the end of Grade 3. Hierarchical linear models indicated that text comprehension at the end of Grade 3 was predicted by spoken word recognition at the beginning of Grade 1. Most importantly, this effect was mediated through lexical quality (phonological and orthographical representations) at the end of Grades 1 and 2.

These results underscore the importance of pre-school language skills, especially the ability to process spoken language, for the development of lexical quality and, as a consequence, of reading comprehension skills.

<i>Johannes Naumann, Yvonne Neeb, Tobias Richter, Julia Knoepke and Maj-Britt Isberner</i>	Process Oriented Assessment of Reading and Listening Skills
--	---

We introduce research aims, methodology, and results of the project „Process-oriented assessment of reading and listening skills in primary school children“. This project has developed a computer-based test battery to assess the efficiency of reading and listening component processes such as phonological recoding, word recognition or semantic integration (“ProDi”, Richter, et al., in press). Efficiency measures are obtained through concurrently measuring accuracy and speed of task completion. These test batteries are currently used to investigate the development of reading and listening component processes during the primary school years in a two-cohort, five-year longitudinal study.

Previous results confirm construct and criterion validity of the reading test battery. In confirmatory factor analyses, the different processes targeted are clearly distinguishable (Richter et al., 2012). Item features that from a psycholinguistic perspective are expected to predict item difficulty, such as the number of orthographic neighbors in word recognition, explain large proportions of item difficulties and response time variability between items (Neeb et al., in preparation). The efficiency of component reading processes, as measured through ProDi, explains substantial proportions of variance in text-level reading comprehension (Richter et al., 2012; 2013).

First results from the longitudinal study indicate that pre-school component listening skills at the word level predict text comprehension skill at the end of grade 3. This effect is mediated through early listening skills at the word level predicting better development of the mental lexicon (lexical quality), which in turn is predictive of better text comprehension skill. Questions we would like to discuss at SWEE relate to the stability of measures across time points and the integration of speed and accuracy into combined test scores.

<i>István Thékes</i>	The Development of a Diagnostic English as a Foreign Language Vocabulary Test Assessing 6th Graders
<p>This study presents our attempt to assess the breadth and depth of English as a foreign language vocabulary (FLV) of 6th grade students. Assessment of vocabulary was carried out with various types of tasks. My goal was to create a test integrating all the skills. I phrased the following research questions: (1) How broad is the mental lexicon of 6th grade students? (2) How do the results scored on the different test elements correlate with one another?</p> <p>When creating the test, I took into consideration validated vocabulary tests of the past 30 years (Meara, 1989; Nation, 1999; Paribakht & Wechse, 1999; Laufer & Nation, 2005). I chose words up to the 2000 band of frequency. The selection of the words were determined on the basis of the frequency lists of the British National Corpus (Kilgarriff, 1997) and professional recommendations as well (Nikolov, 2011). The test included 7 different tasks. Each task contained 9 items plus a distractor. After the finalizing of the test, paper-and-pencil data collection was carried out. 98 6th graders participated in the test in September 2013 in South-Eastern Hungary. The strongest correlation ($r = .82$, $p < 0,05$) was found between task 1 and task 4. No difference was found between students' word recognition whether they heard or read the word. However, there is a significant difference between the results scored on the passive knowledge task types and the active knowledge task types ($t = 2.26$, $p < 0,05$). Looking at the results with a focus on word class, it can also be concluded that the identification of nouns was the most simple for the students and nouns at A1 level were identified by even the lowest scoring students. Classroom implications, limitations and future research are discussed.</p>	

<i>Michel Dorochevsky</i>	Voice and Video capturing in the CBA Item Builder Platform and Implications on Scoring
<p>The CBA (Computer Based Assessment) Item Builder platform developed for the DIPF has recently been extended to support audio and video capturing during the computer assessment process integrated into CBA Items. This functionality is based on the recent Web technology "WebRCT" (Web Real-Time Communications) supported by Google, Mozilla and Opera – submitted as a JavaScript API specification draft to W3C. It opens the field for new types of items where audio or video streams can be collected, in particular for assessing spoken language or presentation skills. These new features will be used in a large scale school study lead by the University of Fribourg/Switzerland in the coming months. These kinds of items cannot be scored automatically based on scoring rules; instead a manual scoring process must be supported where the captured audio or video streams can be scored by a human evaluator. For that purpose we have introduced so called "scoring tasks" in the CBA Items of the CBA Item Builder which provide the same flexibility and power for designing the scoring process as we already have with standard CBA Items. We will present the experiences of the current development and present an outlook on future potential of this technology.</p>	

Session D

<i>Sascha Wüstenberg, Matthias Stadler, Jarkko Hautamäki, & Samuel Greiff</i>	Introduction on Problem Solving and Empirical Results on the Use of Strategies in Complex Problem Solving Tasks
<p>The aim of this presentation is twofold: First, to give a brief introduction on different types of problem solving (i.e., Complex Problem Solving, Collaborative Problem Solving, and Domain-Specific Problem Solving) and to provide an overview of the presentations and round-tables on problem solving from the Luxembourg group. Second, to present recent empirical results on Complex Problem Solving based on a large-scale study including N=3,191 Finnish students. More specifically, it was investigated whether fluid intelligence and students' knowledge of the vary-one-thing-at-a-time (VOTAT) strategy would be sufficient to predict the application of said strategy in an interactive complex problem solving task. Results revealed that fluid intelligence and knowledge of VOTAT are significantly related to CPS strategy. However, a substantial amount of variance in CPS strategy remained unexplained ($\Delta R^2=.583$). Potential explanations such as influence of computer knowledge, metastrategic knowledge, and task effects are discussed.</p>	
<i>Matthias Stadler & Samuel Greiff</i>	The Utility of Complex Problem Solving in University Selection
<p>As the number of young people aspiring higher education rises all over Europe, fair and valid selection of university applicants is becoming more and more important due to a limited number of available places. However, this important challenge has not been met by many universities so far, selecting applicants solely based on their high school grades. Additional predictors complementing established means of applicant selection is needed. Therefore, the aim of this project is to explore the usefulness and feasibility of introducing complex problem solving (CPS) tasks into university applicant selection. Despite several studies providing evidence in favor of associations between CPS and academic success on the level of secondary education, this will be the first study investigating the validity of CPS in predicting university success. In the course of this project we will investigate three related but independent research questions. The first research question concerns the validity of CPS in the prediction of different facets of academic success for university students (A). Secondly, we will explore the validity of additional information gathered from the way participants interact with complex problems (CPS testing process data), on predicting university success (B). Finally, we will research a way to automatically generate CPS items, which will significantly improve feasibility of CPS testing at university entrance and other practical applications (C). Based on current literature, we expect to find CPS explaining incremental variance within university grades. However, due to the flexibility of CPS tasks we will not only limit the definition of success on grades but explore associations between participants' behavior while working on complex problems and real world behavior displayed in the course of their studies. In order to cover the broad range of demands different subjects make on students, we will differentiate between subjects assessing students from business, psychology, and computer sciences. Overall we will aim</p>	

to collect a sample of no less than 200 students. We believe that this project will not only broaden the general understanding of the cognitive skills necessary to successfully achieve a university degree but will actually improve the way universities select their applicants.

<i>Katarina Krkovic & Samuel Greiff</i>	Collaborative Problem Solving: Concept, Assessment and First Results
<p>Collaborative problem solving skills are becoming essential in classroom, at work and in everyday life. For that reason, psychologists and educationalists alike emphasize the emerging importance of measuring and fostering these skills throughout life. However, the assessment of collaborative problem solving is still in its beginnings. In the course of our research, we developed an innovative assessment instrument measuring collaborative problem solving that can be categorized as an experiment-based assessment of behavior sensu Cattell. Complex problem solving tasks are based upon a typical computer-based micro-world approach – MicroDYN - that has been extensively validated for individual problem solving. The instrument uses computer-simulated agents as collaborators and thereby captures the collaborative behavior of students` in a standardized manner. In the first pilot study, we administered the instrument to 92 12-year-old students in one school in Germany. First results are encouraging and show that the instrument functioned as anticipated. Specifically, the results suggest that students manifested stable collaborative problem solving behavior while actively interacting with the computer-agent and working on solving the problem. Additionally, students differed in using specific patterns of collaborative behavior. We discuss the perspective of the assessment instrument, and the necessity of further validation studies.</p>	

Session E

<i>Julia Rudolph & Samuel Greiff</i>	Scientific Inquiry: Conception and Assessment Approach of the Assessment Tool SINQ
<p>Large-scale assessments such as PISA (Programme for International Student Assessment) acknowledge scientific literacy as one of the main competencies that should be fostered in schools all over the world. However, one important facet of scientific literacy is often overlooked in large-scale studies, as its operationalization has only recently become feasible: Scientific Inquiry – the skill to learn laws of nature by self-regulated experimenting. To assess Scientific Inquiry, a computer-based assessment tool is necessary that enables the scoring of strategy behavior. Merging the assessment approach of Complex Problem Solving and static scientific tasks, we have established a computer-based tool named SINQ (Scientific Inquiry) that enables a valid scoring of Scientific Inquiry skills. SINQ exists of 12 small experiments and is designed in a way that is appealing for children and adjusted to the skill level of young students. We will introduce SINQ and its assessment approach and discuss its advantages and shortcomings. In particular, challenges in scoring the student’s strategic behavior will be discussed. Further, first results of a validation study with approx. 350 German students that has been conducted from January to April 2014 will be presented.</p>	

<i>Jakob Mainert, Andre Kretzschmar, Jonas C. Müller, & Samuel Greiff</i>	LLLight´in´Europe 2014 - Update and Political Implication of a State-of-the-Art Large Scale Assessment Project
<p>The practice-oriented large scale assessment (LSA) LLLight´in´Europe provides a constantly growing database of currently more than 550 employees of 19 companies in 6 countries on complex problem solving (CPS), and important demographic, social, economic and educational variables. As an integral part of LLLight´in´Europe, CPS is considered an important transversal skill (i.e., a skill that spans multiple domains) at work with implications for lifelong learning initiatives on an individual, enterprise and policy level. Until the present time, widespread application of CPS with first implications for policy-makers has mainly been restricted to education. The Programme for International Student Assessment (PISA) already links fundamental research on CPS, its reliable and valid methodology and practical applicability for the purpose of policy-oriented outcomes and insights. Surprisingly, similar research on the role of CPS for learning later in life is scarce at best. As an exception, LLLight´in´Europe intends to show that CPS is already measurable in domains of work and has the potential to guide decision-makers in companies and policy. Towards this end, LLLight´in´Europe highlights differences in cognitive and economic performance patterns and identifies inherent features common to high performing employees, enterprises and countries. Apparently, decision-makers are interested in information about how their workforce’s performance compares to that in other enterprises and countries. For example, linking data on employees’ CPS skill level with data on key factors that shape lifelong learning behaviors in and out of work, such as job complexity and the provision of training opportunities, not only provide decision-makers with baseline profiles of skill levels and working conditions in their own and other enterprises and countries, but also allow to delineate activities on how to gain competitive advantage. The goal of this seminar is to provide, discuss and further improve procedures and practice on how to implement transfers from research to policy. The intended outcome is to build know-how to create politically meaningful research results in order to provide interesting and appropriate recommendations to decision-makers.</p>	

<i>Katinka Hardt, Jonas Müller & Samuel Greiff</i>	On the Use and Potential of Process Data
<p>Recent advancements in educational large-scale assessment (LSA) studies involve an increased replacement of paper-and-pencil tests by computer-based assessments. By the implementation of the computer as assessment medium, the collection of data beyond the mere final responses to the tasks and questions posed in the assessment is facilitated. In doing so, every interaction of a student with the computer is captured in so called “process data” (e.g., response times, number of clicks, the buttons pressed and so forth). Besides providing information on behavioral processes during the test-taking process, one major advantage of process data is that they are an implicit, non-reactive measure as students are not aware of their collection. Thus, they allow for drawing inferences on cognitive processes of the students. A disadvantage is that they require the implementation of computer-based assessments, which, in turn, are associated to additional, costly efforts. As LSA studies are an instrument of educational monitoring, several stakeholders are involved. Mainly, they comprise researchers, practitioners, and politicians. After a brief introduction to approaches suitable for analyzing process data,</p>	

costs and benefits of collecting process data within LSA shall be discussed in the light of different stakeholder perspectives. To exemplify ideas and approaches, reference is made to the assessment of complex problem solving, which had been incorporated as a prominent 21st century skill into the most recent cycle of the Programme for International Student Assessment (PISA) study.

Session F

<i>Arthur C. Graesser</i>	Learning and Assessment through Trialogs with Conversational Agents
<p>Learning has occurred over a few millennia by the learner communicating with the teacher, tutor, master, or mentor in natural language. Apprenticeship learning has always occurred one-on-one or in small groups with an expert. Researchers in the discourse and learning sciences have documented the conversation patterns that occur in these interactions. Researchers in computational linguistics, artificial intelligence, and intelligent tutoring systems have developed computer agents that simulate many of these conversation patterns and help people learn. This is the moment in history when these systems are being launched on the internet. This presentation will present recent systems on the internet that help students learn by holding a conversation in natural language. AutoTutor engages in dialogue with the student on a variety of subject matters in Science, Technology, Engineering, and Mathematics. Trialogs are conversations between the human students and two computer agents, such as two peer student agents or a peer agent and a tutor agent. Students can either observe two agents interact vicariously, interact with a tutor agent as a student agent periodically chimes in, or teach a student agent while a tutor rescues a problematic interaction. Agents can argue with each other over issues and ask what the human students think about the argument. Trialogs are being developed for the Internet in serious games with Pearson Education (Operation ARA), in assessments with Educational Testing Service, and in a new Center for the Study of Adult Literacy for struggling adult readers. AutoMentor is being developed for computer mediated communication between a mentor agent and small groups of students in a simulation game on urban planning.</p>	

<i>Zsolt Lavicza</i>	Integrating Technology into K12 School Teaching to Enhance STEM Education in Hungary
<p>Technology is increasingly becoming an important part of STEM teaching and learning in the 21st Century. There have been numerous attempts to integrate technology into education systems, but without serious development and research the success of these attempts have been limited. Recently, together with my colleagues in Hungary, we started a project called GeoMaTech to develop high-quality teaching and learning materials for all grades in primary and secondary schools in Hungary. These materials (1200+ Mathematics, 600+ Science) will be embedded into an on-line communication and collaboration environment that can be used as an electronic textbook, a homework system, and a virtual classroom environment. In addition to material development, we will offer 60-hour professional development courses for more than 2500 teachers in 800 schools in Hungary. Furthermore, we will organize a wide-range of teacher and student activities including competitions, maths and science fairs, and develop a network of schools for the long-term</p>	

sustainability of the GeoMaTech project. The technology background of the project is offered by GeoGebra (<http://geogebra.org>) , which an open-source, dynamic mathematics software used by more than 30 million students and teachers around the world. All activities of the GeoMaTech project will be assisted and evaluated by a strong research team offering support for pedagogical resources involving teaching methods based on highly-respected Hungarian Teaching Traditions by Pólya, Lakatos, Varga and Dienes as well as successful technology integration programmes from other countries. The research team will also carry out a pilot programme with approximately 3,000 students and prepare instruments for evaluating the impact of GeoMaTech on teachers' and students' learning and understanding of mathematics and science. In addition, in our project we aim to involve and collaborate with as many experts and researchers as possible from around the world. We plan to base our work on previously successful projects, collect ideas from the GeoGebra community, and invite colleagues to work with us in Hungary. We hope that this project could be a test bed for future projects and trialling ground for different ideas. In my talk, I will outline the GeoMaTech material development and teacher training initiatives, building a network of participating schools, the directions for software development, the aims of the GeoMaTech research and evaluation team, and hope we can discuss possible collaboration opportunities.

<i>Carolin Hahnel, Frank Goldhammer, Johannes Naumann and Ulf Kröhne</i>	Step by Step: Explaining Relations between ICT Skills and Reading Digital Texts by Navigation
--	---

Digital media has changed the way text is presented and read. Especially when reading on the Internet, students need skills to deal with Information and Communication Technologies (ICT) in addition to their text comprehension skills. In this study, we argue that relations between ICT-related skills and digital reading are linked by students' navigation behavior. Depending on the conceptualization of navigation, we expect navigation to mediate between ICT skills and digital reading in a specific manner. Considering navigation as global moving through hypertexts, students who have lower basic computer skills might navigate too less or too much, respectively, resulting in less effective digital reading. Higher skills in dealing with computer interfaces, however, should be associated with moderate amount of global navigation steps resulting in better digital reading performance. If navigation is conceptualized as task-oriented process, students who are more able to evaluate the utility of an encountered online source efficiently are supposed to visit more pages that are relevant for their reading purpose. As a consequence they should show a deeper understanding when reading digital texts. In the context of PISA 2012, we assessed a total of 888 15-year-old German students (46.6% female) and tested their skills in digital reading, basic computer skills, evaluation skills on online sources, and print reading. To determine navigation indices, data from students' log-files in digital reading was extracted. As indicator of global navigation, the total number of page visits was used. The proportion of total visits on relevant pages to total visits on all pages represents task-oriented navigation. First results show that in addition to reading comprehension ICT skills can explain unique proportion in variance of digital reading, supporting the assumption about the importance of ICT skills in digital reading. Mediation analyses from basic ICT component skills to digital reading through students' navigation, however, are not processed yet.

<i>Heiko Rölke, Ingo Barkow</i>	TIPO – TAO-based Item Portal
<p>Over the last years, more and more studies have been performed using computer-based assessment instead of paper, even on the international level. Software support for single studies is available and generally usable. In longitudinal studies, however, additional requirements arise for managing users, items and tests. We show how the TAO software can be extended and supplemented with extra functionality to fulfill these requirements without compromising future compatibility with new TAO versions. While offering a quite good coverage of the overall assessment cycle, certain features important for bigger contexts are still missing in TAO. Examples are extended item search and filtering, item versioning, restricted access to the back-office, and roles and rights management. TIPO is a portal solution acting as a façade to the TAO system. It restricts the usage of the features offered by TAO according to the role of the user. In addition, it offers tailored access to certain features generally available in TAO, but sometime difficult to use. All information is stored as metadata in TAO. Furthermore, TIPO integrates additional item types to TAO, e.g. the CBA ItemBuilder. Typical use cases for TIPO are large-scale, longitudinal assessment projects with many users in different roles over longer periods of time. The development is currently done in cooperation with the Swiss federal government.</p>	



The project is supported by
the European Union and co-financed
by the European Social Fund.