University of Szeged
Faculty of Arts
Doctoral School of Education

ATILIA RAUSCH

ONLINE ASSESSMENT OF EARLY NUMERACY

Summary of the PhD dissertation

Supervisor:
Prof. dr. Benő Csapó

2018
Szeged
SUBJECT OF THE DISSERTATION

Mathematics has its special place among sciences. However, knowledge and application of basic mathematical procedures are not only required for science, basic mathematics is needed in various situations of everyday life as well. Moreover, there are less and less jobs without the necessity of some mathematical knowledge, or thinking (Nunes & Csapó, 2011). Meanwhile, the results of the PISA assessments showed that there are serious problems with the mathematical literacy of Hungarian students, with the decrease of average achievement, the rate of drawback students has increased dangerously (Csapó, Fejes, Kinyó, & Tóth, 2014). Huge number of students cannot apply the mathematical content they learned in the elementary level of school. Researches of the last decade have shown that the origins of the problems can be found already in the time of school entry, with increasing attention being paid to family background, selection and the importance of early mathematics education. The level of early numerical skills during kindergarten to school transition influences students' later mathematics performance (Jordan, Kaplan, Ramineni, & Locuniak, 2009). Early numerical skills comprised of symbolic and non-symbolic components, which are connected with skills and knowledge that are essential to learn mathematics in school, and early numeracy plays an important role understanding the mayor mathematical content areas. Numerical skills include the knowledge of names, order, and sign of numbers, recognition of magnitudes, and the ability to perform numerical operations, in addition to understand basic mathematical concepts. Drawback in early numeracy may contribute some interruptions understanding new mathematical concepts and block the progress of learning mathematics in school. Therefore, developmental programs play important role in early childhood education. According to previous studies the level of basic skills can be greatly influenced in kindergarten age (Józsa, 2004), and efficiency of early interventions has been proven in several cases (e.g., Molnár, 2011). In order to provide development for children with drawback, we need to identity problems in time, and this is not achievable without assessment tools that can be used instantly and easily for all teachers, while providing an accurate and reliable picture of the level of early numeracy skills of children. Nowadays, the development of digital technology offers several advantages, especially for educational assessment and evaluation. The use of technology-based assessment for educational purposes was unfolding in the past decades, and its potential can be used for younger children as well (Csapó, Molnár, & Nagy, 2014). Online testing of early numerical skills is become reachable. Although this form of assessment has many opportunities, the range of challenges is also high. The spread of technology-based assessment could help to identify pupils with drawbacks who need more attention, and with electronic testing we are able to save time to focus
on fostering their skills, which is also more efficient using results of prior assessments. In our research, we aim to develop a new online test to assess early numeracy. We present the process of the test development, in addition to the results and discussion of our assessments in kindergarten and primary school.

THEORETICAL BACKGROUND

Early numeracy is an important basis of mathematics learning, it plays role in the development of all areas of mathematics. Various models have emerged in the recent decades to describe early numerical development in early childhood. Some models focus on the simultaneous development and interconnection of different numerical skills (Aunio & Räsänen, 2016), while others build on the construction of elements of early mathematical knowledge and experiences (Fritz, Ehlert, & Balzer, 2013).

Understanding the development of early numeracy, Dehaene's (2003) theory can be considered as a milestone that has since been studied and has been supplemented and improved (Igacs, Janacsek & Krajcsi, 2008). It includes not only the three main components, the quantitative, verbal, and the Arabic numerical system, but also the interconnection of the components that play an important role in numerical development, in addition to mathematical concepts, procedures and knowledge. Symbolic numeral conceptualization starts by learning the name of numbers. The verbal number word sequence skills are an important basis of other early mathematical skills. The knowledge of the correct order (forward or backward) of number words is essential to the development of enumeration skills, and it has significant role in solving basic additions and subtractions (Józsa, 2014). Enumeration is also an important component of the basic counting skills, it is related to the cardinal meaning of numbers when children identify the last number of the sequence with the number of the element they counted (Aunio & Niemivirta, 2010). Basic counting skills develop swiftly after children are aware of number word sequences and understand the cardinal number concept. They are able to solve additions, subtractions and later they understand the part-part-whole concept as well (Fritz, Ehlert, & Balzer, 2013). Another essential component of early numerical skills is the knowledge of number symbols. Numeral recognition and number identification are the two segments of this factor. At the beginning children learn Arabic numerals then they will be able to identify and read numbers. Establishing connections between Arabic symbols and quantities are also important basic skills of numeracy.

Studies have shown early numeracy as a strong predictor of later mathematical achievement (Jordan et al., 2009; Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Friso-
van den Bos, Kroesbergen, & van Luit, 2014). There are several standardized test batteries to measure the numerical skills of children from age four to eight, we reviewed four of them in our theoretical background. A widespread international assessment tool is the ENT (Utrecht Early Numeracy Test) (Van Luit, Van de Rijt, & Pennings, 1994), which were used in several studies all across Europe, another well-known test which is mostly used by teachers called a WENT (Wright Early Numeracy Test) (Wright, Martland, & Stafford, 2006), in addition to MARKO-D (Mathematik und Rechnen Konzepten – Diagnose), a diagnostic assessment tool developed by German scholars (Ricken, Fritz, & Balzer, 2013). In Hungary, a diagnostic test battery called DIFER (Diagnostic System for Assessing Development for four- to eight-year-old children) is widely used to assess key skills for school readiness, including basic counting skills (Nagy, Józsa, Vidákovich, & Fazekasné Fenyvesi, 2004). These instruments share the same characteristics: they are reliable, however they require resource and time consuming face-to-face test administration and the educators’ proper qualification is also necessary otherwise the objectivity of the measurement can be compromised (Csíkos, András, Rausch, & Shvarts, 2018).

The potential of technology-based assessment can also be used among young children. Pre-recorded instructions can help to overcome reading difficulties, simulations, colorful illustrations, and manipulative tasks are suitable for the needs of young children (Molnár, 2011; Rausch, 2016b). Recently, more and more researches have been carried out to explore the possibilities of technology-based assessment of young children (Csapó & Pásztor, 2015), including basic counting skills as well (Csapó, Molnár, & Nagy, 2014). Exploring the possibilities of online assessments in kindergartens has only started in the last couple of years (Csapó, Hódi, Kiss, Pásztor, Rausch, & Molnár, 2017). By overcoming the possible difficulties of using ICT tools, measuring basic skills of young children is widely approachable. Csapó, Molnár and Nagy (2014) studied the possibilities of computer-based testing of school readiness among first-grade students. They compared the results of the original and the online test version of the DIFER. They highlighted that in some areas technology-based assessment has special technical advantages, and students’ performance (e.g. phonological awareness) can be more reliably evaluated, but they also shown that there are areas where it is not possible to properly transform tasks to an electronic platform. In their study, the DIFER’s basic counting skills test was among them. All of the test items could not be tested with computer, and the computer-based assessment was less reliable than the original face-to-face test. Therefore, it is necessary to develop new tests or test items specially to the new digital assessment platforms while taking into account the technical possibilities, the way of responding, and the interface (Pásztor, 2017; Rausch & Pásztor, 2017).
AIMS OF THE RESEARCH

The purpose of our research is to develop an online test to assess children's early numeracy from age 5 to 7. In the dissertation, we present an online assessment tool that quickly and reliably measures the development of early numeracy of young children. In our work, we present the details of the test development with the psychometric characteristics of the test.

Besides the test development, we also aimed to examine the development of early numeracy skills and their impact on early mathematics education. Assessments were made in two age groups, in kindergarten and in primary school with the same test, therefore we can perform cross-sectional comparative analyses. In addition, we also had the possibility of tracking the students in first grade, our longitudinal analysis provides information about the connection between the level of early numeracy at school entry and mathematics achievement at the end of the first year of school.

METHOD

Sample

The studies were carried out among children from kindergartens and primary schools. The first two kindergarten assessments were carried out in Szeged, with the management of the Research Center of Learning and Instruction, University of Szeged. Overall 342 children participated in the first kindergarten measurement, their average age was 5.8 years, 252 children were examined on the second measurement, the mean of age was 6.4 years. Sample of the validation study comprised of 30 children, their mean of age was 5.7 years. Sampling of the kindergarten measurements were determined by the reachability of the institutions. The school assessments are part of the Longitudinal Research Program of the MTA-SZTE Research Group on the Development of Competencies, the tests were administered in the 4th sample of the program (Csapó, 2014), which was based on nationwide territorial representative sampling. In this school entrance test, we studied 5154 first-year students in the autumn, and the spring a second measurement was carried out, when 4277 pupils were measured. The mean age of the sample was 7.1 years at the first point of measurement, and 7.7 years for the second measurement.
Instruments

We started to develop our own online early numeracy test with considering several scientific and technical aspects of technology-based assessment and early childhood assessments. We aim to cover numeric skills that needed to learn mathematics in school, and our major goal was to develop an online test with automatic evaluation and instant feedback. Moreover, we paid special attention to the time of testing, the quantity and design of the tasks. The main components, content and difficulty of our test were optimized for kindergarten and school entry level, so for children from age 5 to 7.

The test had two versions, the first version had six subtests, and after a revision we created a new, second version with four subtest. The basic counting subtest contains manipulative tasks, addition, and subtraction of magnitudes, and tasks related to the part-part-whole concept. Within these tasks children need to add, take away or sort the right amount of magnitudes. The number word sequence subtest measures whether children can recognize a correct forward or backward number word sequence. They hear a sequence of three numbers, then they can listen to three possible conclusions and they need to decide which is the correct one. In the numeral recognition subtest children need to recognize Arabic numbers with one, two and three digits. They select the right card out of four that shows the number what they hear. In the magnitudes and numerals subtest children manipulate magnitudes based on the number they hear or see. Tasks with smaller amounts are solved by drag and drop technique but tasks with larger amounts can be solved by the selection of the right picture of three different magnitudes. Tasks of the relations subtest measured whether children can compare number sets and find the larger, largest or smaller, smallest quantities. The test was developed and the assessments were administered through the eDia online assessment system (Csapó & Molnár, 2017; Molnár, 201).

Besides our online early numeracy test, we used different measurement tools in our studies. To validate our online assessment, we used two tests, the counting and basic numeracy, and the relational reasoning test of the Hungarian DIFER test battery (Nagy et al., 2004). The counting and basic numeracy the test was reliable (Cronbach-α=.80). The relational reasoning test which assesses the understanding of words which stand for relations between different objects, attributes or processes, had acceptable reliability (Cronbach-α=0,62). First grade students were participated in a second assessment with a mathematics achievement test at the end of the first school year. Tasks of the mathematic test were selected from an item bank created within the Developing Diagnostic Assessments research program (Csíkos, Molnár, & Csapó, 2015). The whole test and its three dimensions (thinking, application, discipline) were reliable, and the overall test reliability was excellent (Cronbach-α=.94.), reliability of the subtest were
good (Cronbach-α= .83-.88). All of our measurements included an ICT literacy test as well (Molnár & Pásztor, 2015). During online assessment, especially in early childhood it is important to check whether students are having any problems using devices such as tablets or personal computers to solve the tasks. This tool comprises of tasks, item types which are similar to the early numeracy test items, and during this test they are practicing as they are being measured. Reliability of the ICT literacy test is different within the studies (Cronbach-α=.40-.72). It can be caused by the ceiling effect, children performed well on this test in each measurement (Mean=90.3-95.1 %p; σ=7.5-13.4 %p).

Procedure

The children solved the tasks of the tests independently in all of the online assessments, assistants only intervened when technical difficulties raised, and they helped to start the testing process. In addition, our online tests were preceded by a ICT literacy test (Molnár & Pásztor, 2015), the test was presented previously within our instruments. Therefore, students could meet and practice solving tasks on electronic platforms. Kindergarten assessment took place in a separate room of the kindergartens. School measurements were carried out in computer laboratories of the participating schools and assessments were monitored by class teachers of the schools. They received a detailed measurement instruction for the proper conditions of data collection. Within all of our research, we kept ethical rules of social sciences, covering the planning and implementing phase, and the communication of those affected by the measurements, ensuring anonymity, as well as data management and analysis.

RESULTS OF OUR RESEARCH

Online assessment of early numeracy in kindergarten

Reliability of the early numeracy test in the first kindergarten assessment was high (Cronbach-α=.94), all subtests were also reliable, their values of Cronbach-α were between .70 and .88. The number word sequence subtest, which was based on dichotomous choice, had the lowest reliability. Items of the test were also examined by Rasch analysis, the EAP/PV reliability of the test was .92. The average performance on the test in kindergarten was 57.2%p, with standard deviation of 23.3%p. The distribution shifts to the right, and this shift is not symmetrical. Additionally, the Rasch analysis showed that primarily the lower ability levels were covered by our tasks, and there are
no tasks in the test between ability levels 1 and 2. To verify the construct validity of the early numeracy test, a confirmatory factor analysis was performed. The 6-dimensional CFA model is describing the structure of the test well ($\chi^2=190.82$, $p<.001$, $CFI=.953$, $RMSEA=.043$). The standardized factor coefficients were high for each of the subtests. The structural validity of our first test variant can be considered good at the first kindergarten study.

After a revision, we created the second test version of early numeracy test, with a total of 34 items and four subtests, and we used this version in the further kindergarten studies. The reliability of second version of early numeracy test was lower compared to the previous kindergarten measurement but was still acceptable ($\text{Cronbach-}\alpha=.83$). Based on the results of Rasch analysis, the test EAP/PV reliability was .83, which dropped but remained acceptable. Construct validity of the revised test was decent, the model fit of the 4-dimension CFA-model is acceptable ($\chi^2=1266.8$; $p<.01$; $CFI=.907$; $RMSEA=.06$). The test items has high coefficients toward their scales, and the four subtests are strongly connected to the main factor of early numeracy. The average performance of kindergarten children was 62.9%. According to the Rasch analysis, the items grouping around ability level -1, and 2, and between this two difficulty level there are only a few items. Based on the results, further development of the test is still necessary, so we can cover the children's ability levels more evenly (Rausch; 2016a; 2017c).

**Online and face-to-face assessment of early numerical skills**

Based on the results of the presented studies, we have been able to explore the reliability, structural and predictive validity of the online test. But we still have little information to find out how well it can replace face-to-face measurements. For this purpose, a small sample kindergarten study was carried out (Rausch & Pásztor, 2017), and besides two online tests, early numeracy and ICT literacy (Molnár & Pásztor, 2015), the DIFER’s basic counting skills and relational reasoning tests were also administered (Nagy et al., 2004). Results show that the performances of the online and face-to-face test are related, and we found correlation between the two numerical test. There is a significant correlation between the online early numeracy test and the DIFER basic counting test ($r=.84$; $p<.01$). We found significant correlation between the online number word sequence subtest and the DIFER’s number word sequence subtest ($r=.49$; $p<.01$), and the DIFER reading numbers and the recognition of Arabic numerals in online form were also correlated ($r=.76$; $p<.01$). Another result of the study is that we found connection between the face-to-face relational reasoning test and the ICT literacy test ($r=.48$; $p<.01$). During the previous measurements, we found significant correlations between
the ICT literacy and early numeracy, however ceiling effect was measured on the ICT test in all cases. Therefore, using tablets and computers and solving task on these digital platforms were not difficult for the young children. Based on the relationship with the relational reasoning, it is assumed that the ICT literacy test can measure not only operations of using tablets or computers, but also a general task solving or speech comprehension, which can be explored in the future. This validation study is based on a small sample, however the results show that online assessment can provide a good estimation of the development of early numeracy such as the face-to-face measurements (Rausch & Pásztor, 2017).

*Online assessment of early numeracy at school entry*

Possibilities of the online assessment of early numeracy was also examined with a primary school entry assessment (Rausch, 2017b; Csapó, Rausch, & Pásztor, 2016). Research study among first grade students with the test shared the same psychometric characteristics as the kindergarten assessments. The reliability of the test was lower, but it was still acceptable (Cronbach-α=.89). The subtest were also reliable Cronbach-α values lied around 0.7. Additionally, the EAP/PV reliability index of the full test was .86 based on the Rasch analysis, which can be considered acceptable. The structural validity of the first test version of the early numeracy test was also examined by this study in primary school, and the fit of the CFA model with the six subtest was good ($\chi^2=5089.56; p<.001; CFI=9.28; RMSEA=.034$). Values of CFI and TLI are also acceptable, with minimal deviations between them, and RMSEA level of .03 is considered to be excellent. Moreover, we were able to have look on the predictive validity of the test, thanks to the results of the first grade mathematics test. The values of the correlation coefficients between test items of early numeracy test and the result of the mathematics achievement test are significant in all cases. The easier tasks at the beginning of the subtests had lower predictive validity ($r=.13-.19$), and the tasks at the end of the subtests had higher coefficients ($r=.19$ to .42).

With the first test version of the early numeracy test, we performed assessment in kindergarten and in primary school as well, so we have the opportunity to conduct cross-sectional comparison between the two age groups to show the development of each component. The level of development from kindergarten to first grade is significant (Cohen-d=1.05). The performance of the early numeracy test was 21 percentage points higher among primary school children. The average performance of all subtest increased by 20-21 %p. The first grade students reached 80.5 %p ($\sigma=23.5 %p$) on the early numeracy test. Distributions shifted strongly to the right, and the Rasch analysis showed that most of the tasks were too easy in this age. Higher ability levels, apart from one or
two exceptions, are completely missing from the test. So, the test was able to measure only one half of the sample, and a large part of the students did not meet with task fit to their level of skills (Rausch, 2016c; Rausch, 2017b).

Effect of early numeracy on learning mathematics in primary school

Longitudinal analyzes had also been carried out with the first version of the early numeracy test. First-grade students participating in the school entrance measurements also participated in an assessment at the end of the academic year, when they completed a mathematics achievement test. According to the results of the linear regression analysis, the level of early numeracy at the start of school significantly interprets the mathematics achievement (28.4 %). For a more detailed examination of the effect of early numeracy on school mathematics, we applied structural equation modeling (SEM), which was used to examine the model fit of three hypothetical models. In our first model the effect of the early numeracy on the mathematics results was examined. The model fitted well ($\chi^2=363.4; p<.001; \text{CFI}=.972; \text{RMSEA}=.055$), and it shows that early numeracy has a significant effect on mathematical performance, the standardized regression coefficient between the two major factors is .63. In the next model, the effect of early numeracy skills was examined separately for the results of the three mathematics subtest. Fit of this model was also good ($\chi^2=355.7; p<.001; \text{CFI}=.973; \text{RMSEA}=.057$). We found significant regression coefficients on thinking, .48, on the application, .51, and on the discipline dimension, .55 ($p<.05$). In our last model, we analyzed the effects of the components, based on the six subtests of early numerical skills on mathematical performance. Although the effects of each subtest were significant, only the relations (.37) basic counting (.13), and the recognition of the Arabic numerals (.11), had the regression coefficients above .1. The model fit was excellent ($\chi^2=37.48; p<.001; \text{CFI}=.996; \text{RMSEA}=.022$). Our findings support the results of several previous researches, that the development of early numerical skills has effect on later mathematics achievement (Aunio & Niemivirta, 2010; Jordan et al. 2009), results of the study also highlight the role of early numerical skills in early mathematics education (Rausch, 2017a).

Educational relevance

The main product of our research is a reliable and valid online assessment tool that can be used widely by kindergarten and class teachers to assess the early numerical skills of their students. They will be able to screen easily children who may have drawback in
the development of early numeracy. Based on the results of online measurements, they can design their early mathematics education and training sessions. With the potential of online measurements, we can provide immediate feedback on individual and group level as well, and recommendations for the content and methods of development automatically can be formed in the future. We also want to support the work of teachers by self-developed development programs which are connected to the results of our early numeracy test (Rausch, 2015; Rausch, Debreczeni, & Szabó, 2014; Rausch & Turainé Toldi, 2016).
REFERENCES
(WITHOUT SELF CITATIONS)


CITED PUBLICATIONS OF THE PHD CANDIDATE


