UNIVERSITY OF SZEGED
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
DOCTORAL SCHOOL OF EDUCATION
INFORMATION AND COMMUNICATION TECHNOLOGIES IN EDUCATION

Alisa Tóth

ONLINE ASSESSMENT METHODS AND RESULTS OF ASSESSMENT OF COLOUR PERCEPTION AND INTERPRETATION AMONG 6-TO 13-YEAR-OLD CHILDREN

Dissertation Theses

SUPERVISORS: PROF. DR. ANDREA KÁRPÁTI
PROF. DR. GYÖNGYVÉR MOLNÁR

Szeged, 2020

THEME OF THE DISSERTATION

The Information Society of the 21th century, an epoch also called the Age of the Image, is facing ongoing challenges in education: it requires the acquisition of even more complex skills and abilities, including creativity, critical thinking, and well-established visual communication techniques that is mainly developed through disciplines related to art and design education (Bintz, 2016).

An important aspect of knowledge lies in its applicability in other, new situations, including recognizing and interpreting the connections between knowledge elements (Molnár, 2002). There is a strong relationship between the concept of knowledge transfer and the findings of a new model of visual literacy, described by the Common European Visual Literacy Framework. ”Situations” are arts education tasks contextualized in real life that activate competencies in an authentic setting. (Kárpáti & Schönau, 2019; Pásztor, Babály, Simon, & Tóth, 2017; Wagner & Schönau, 2016). The presence of visual competencies is necessary for many workplace situations, where fast, creative and efficient knowledge is required in communication and interpretation ideas.

Colour belongs to the elements of design such as line, shape, or dot (Bakos, Bálványos, Preisinger, & Sándor, 2000). From all visual art elements, colour has the most powerful communicational feature. Despite the fact that international and Hungarian arts education curricula prescribe colour trelated skills and knowledge in detail for all age groups, (Ministry of Human Resources, 2012a), research results on the development of colour perception and colour interpretation are largely unavailable and an elaborated description of colour perception and interpretation for developmental program design also seems to be lacking.

Understanding special effects of colour on psychological functioning is important for designing and shaping our immediate environment. It enriches our visual culture, helping to understand, interpret and engage with works of art that also contribute to our mental well-being. For a correct interpretation of visual signs that have become major orientation codes in many everyday situations, their perception and interpretation has to be acquired. In the 20th century, research on visual skills and abilities focused on mapping the skills system and defining its developmenta levels and offered artistically authentic tasks for skills development and educational assessment (Schönau, 2012).

Research presented here offers a conceptual model of colour perception and interpretation based on the Common European rFramework of Visual Literacy, utilising research results and curriculum analyses from 21 EU countries. The dissertation focuses on issues of test development processes of colour perception. It offers a testing tool for colour interpretation assessment that was piloted and analyses the factor structure of the underlying theoretical model. The task development process and resulting data were analysed both with classical test theory and item response theory. The construct validity and its dimensionality are explored with techniques for building structural equation as well.

THEORETICAL BACKGROUND

Online assessment of visual skills

The evaluation of the system of visual skills is based on two competence clusters: creation and interpretation. Our assessment efforts target skills and abilities developed by Visual Culture, the Hungarian discipline for arts education. Our skills model also takes into consideration the assessment of psychological functioning of visual skills and abilities, that are researched independently from school subjects.

Between 2009 to 2011, the Hungarian Framework of Visual Literacy was developed (Kárpáti & Gaul, 2011). This project was implemented in the program "Development of Diagnostic Assessment" Project, coordinated by the Center for Research on Learning and Instruction, University of Szeged (Kárpáti & Gaul, 2011). During the second phase of the project, from 2011 to 2013, two of these skill clusters were assessed: spatial s visual communication skills (Kárpáti, Babály, & Simon, 2015). The empirical findings of the second research phase served as a basis for further research aiming to map and develop (1) visual communication (Kárpáti & Simon, 2014; Simon, 2015) (2) spatial perception (Babály, Budai, & Kárpáti, 2013; Babály & Kárpáti, 2016) and (3) perception, interpretation and creation of colours (Tóth, 2017; Tóth, Kárpáti, & Molnár, 2017; Tóth, Molnár, & Kárpáti, 2019).

The theoretical framework of assessment of colour perception and interpretation

The previous chapter presented research progress on assessment methods related to the Hungarian Framework Visual Skills and Abilities. The theoretical framework presented in this dissertation is connected to the skills and abilities of the first and fourth cluster of visual skills identified in the Hungarian Framework of Visual Literacy: visual perception and visual communication. As research on the influence of colour on the visual skills development of children is scarce, it is hard to implement the empirical findings in visual arts education research and practice.

The conceptual framework of colour perception and interpretation is based on the findings of colour perception assessed in other domains, which could be defined along four subskills: colour sensitivity, colour-and form recognition, colour memory and colour and meaning.

Colour sensitivity

Perception of colours considers three attributes: hue, chroma and lightness (Elliot & Maier, 2014). Hue separates colours from each other, chroma refers to the vividness of colour, lightness is the white-to-black equity of the colour. According to Abramov et al. (2012), women have no problems in distinguishing subtle shades of colours, whereas men are more likely to have, but research reports that males have significantly better sensitivity for fine details and for rapidly moving stimuli perceived in space. They explain their findings through partially evolutional, partially biological differences, but suggest further research (Abramov et al, 2012).

Colour-and form recognition

Colour leads to better performance in shape-and texture recognition. This finding is empirically supported by experiment with pictures of natural scenes performed by Gegenfurtner & Rieger (2000). McGivern, Huston, Byrd, King, Siegle & Reilly (1997) observed significantly better performance in females for recognition of visual stimuli, then males. Moreover, they explored age-related differences among teenagers: 10-to 11-year old children achieved one or three fewer correct responses than 15 year olds. Gathers, Bhatt, Corbly, Farley, & Joseph (2004) studied neural networks of face and object recognition, emphasizing the importance of these kinds of learning processes and experiences in cognitive development. Confirming the developmental hypotheses, they found that facial recognition strategies for nine-to eleven-year olds are more similar to those of adults than younger mates (Gathers et al., 2004).

Colour memory

Research on colour memory is conducted from the aspect of different disciplines, including experimental psychology. Findings suggest that explicit colour memory show age-related differences (Mecklenbräuker, Hupbach & Wippich, 2003). They revealed that 14-year-olds achieved significantly better results in explicit memory tasks than younger and preschool aged participants. To clarify this statement, Mecklenbräuker and her colleagues changed the test instruction for the explicit memory during the test phase, where they instructed pupils to observe the colours during the priming.

Colour and meaning

When perceiving colour, we are also aware of the content of visual information (Albers & Harste, 2007). Because of its strong emotional influence, spearating the assessment of colour and meaning is challenging. However, by focusing on the informative meaning of colours, we can learn a lot about about the ways of their interpretation. A growing body of research reports on children's meaning making processes related to colours. Willson, Falcon & Martinez (2014) explored children's visual language through activities of creation and interpretation. They confirmed Arnheim's statement (1986) that is consistent with Sipe's theory: experience gained through images is the sensory equivalent of understanding meaning on the cognitive level (Sipe, 2008; Willson, Falcon & Martinez, 2014). Visual perception is not automatic, it grows through experiences and palys a major role in the acquisition of visual language. Children's development level related to colour and meaning are expanded through creative activites (Burkitt, 2004; Zentner, 2001; Cox, 2005; Siu, Lam, & Wong, 2015).

Siu, Lam & Wong (2015) reported gender differences in colour and meaning with a creation tasks among elementary school children. They revealed sex related differences in favour of girls. Hungarian studies on visual skills and abilities confirmed that interpretation of colours develops with age (Kárpáti, 1996; Gaul-Ács & Kárpáti, 2018). Gaul-Ács & Kárpáti (2018) studied the development of visual skills of kindergartener children though complex, creative tasks. Assessment items included the creation of coloured drawings. The resulting data confirmed age and sex related differences: interpretation of colours develops with age, girls

outperformed on test and subtest level - a finding consistent with results revealed by Siu, Lam & Wong (2015).

Teaching about colours: curriculum analysis, major conclusions

In the National Core Curriculum, knowledge about and creation with colour is an important knowledge field (Alaptanterv (Ministry of Human Capacities, 2012a, 2012b, Ministry of Education and Training, 2010). Requirements include the interpretation of essential colour compositions in fine arts, media and applied arts, and interpretation of functional and expressive uses of colour signs and symbols.

This chapter analyses expectations, knowledge, skills and abilities related to teaching colour in the curricula of Anglophone countries: Ontario, British Columbia, United Kingdom, Australia and USA, and Hungarian curricula. Considering the curricula analysis, we found that colour appears very often in the curricula of Anglophone countries, often linked to the immediate environment (e.g. school, nature). Development of skills and abilities related to colour perception are emphasized in all documents. English, Australian, and American art education focus on creation with colour and dedicate less emphasis on teaching colour and applying colour knowledge in lessons than in Canada. Compared to other curricula, Canadian visual arts education encourages more complex thinking strategies. The in-depth study of visual language is in contrast to traditional perceptions of arts education preferring „pretty little drawings" (Kárpáti, 2011).

FRAMEWORK OF COLOUR PERCEPTION AND INTERPRETATION POSSIBILITIES OF ONLINE TEST ENVIRONMENT: RESEARCH AIMS, HYPOTHESES

Research aims of the dissertation are the following:

1) Development of an online task bank for assessing colour perception and interpretation skills in a wide age range, which provides immediate, automatic and personalized feedback for art teachers about results of their educational efforts in this area.

2) Empirical validation of the four-dimensional, hypothetical model of colour perception and colour interpretation based on findings of literature analysis.

3) Identification of the developmental features of colour perception and colour interpretation, examination of its structure, and description of the differences of performance of elementary school students of lower and higher grades.

4) Identification of gender-related differences in achievement.

5) Identification of background variables influencing students'skills development level, exploring factors predicting development.

Hypotheses (H) are systemized in accordance with the research aims above:

H1: It is possible to develop a reliable online test item bank, which assesses colour perception and interpretation skills to be used for testing students in lower and higher elementary grades.
H2: The conceptual framework is valid (construct validity), therefore, the empirical data fit to a four-dimensional theoretical model (construct validity). Components of the model: colour sensitivity, colour-and shape recognition, colour memory, and colour and meaning
H3: Based on findings of curricula analysis, we assume that there is a significant improvement during primary school years, there are significant differences between the lower and higher primary school children the colour perception and interpretation skills and abilities, explored on dimensions-level. We assume that the structure of the construct is changing over time s observed in other visual domains (Gaul-Ács & Kárpáti, 2018; Kárpáti, 1996).
H4: According to our assumption, there will be significant differences in colour perception and interpretation development level among boys and girls and in both age groups as observed in Kindergarden age children (Gaul-Ács & Kárpáti, 2018).
H5: According to our hypothesis, the evaluation of the Hungarian discipline for art and design education Visual Culture, gender of the students and their attitude towards art have a decisive influence on development level of the assessed ability.

The research presented in the dissertation is based on two previous research efforts that involved definitions of colour perception and interpretation: the Hungarian Visual Skills Framework and the Common European Framework of Visual Literacy (Kárpáti & Schönau, 2019; Wagner & Schönau, 2016).

DEVELOPMENT OF AN ONLINE TEST FOR ASSESSMENT OF COLOUR PERCEPTION AND INTERPRETATION DEVELOPMENT: PILOT RESEARCH

Assessment of colour perception and interpretation development level among lower primary school children: pilot studies

The previous chapters of the dissertation introduced the theoretical model for the assessment of colour perception and interpretation and outlined the major findings of international visual arts curricula in relation to teaching about colours.

The elaboration of tasks feasible for assessment of the construct was a result of a longer development process involving several pilot measurements. In the first steps, we constructed online tests for first and second graders. Further characterization of the behaviour of the test of pilot data acquisition is focused on analyses of scales and subscales of the sample. Large-scale assessments were prepared through the second small sample data collection which aimed to compare the achievement of creation and perception of colours among second-grade students in traditional and online environments. The objective of comparative analyses was to justify the

dual property of visual skills and ability system through creation and perception, whereas in Hungarian arts education practice they are divided activities in the system of visual skills and abilities.

The online test assessed perceptual skills, while the traditional instrument assessed creative skills. Whereas one of the research objectives was to reveal developmental trends of colour perception and interpretation, during the test developmental processes from the test battery constructed for second-grade students we selected the most difficult items. These selected difficult items were applied as anchor items and expanded with more difficult, new items feasible for age groups in higher elementary school. Before large sample data collection among seventh-grade students, a newly developed online test was carried out among five to seventh-grade students. The next chapters of the dissertation will present the process and difficulties and challenges of test development; the procedures of pilot studies, the results and the drawn conclusions related to pilot studies.

Data of the small size samples were analysed both with classical test theory (e.g.: applying classical test theory indicators to reveal the different characteristics, e.g. gender-related differences), item response theory (e.g. one-and multidimensional person/item maps) and with techniques for building structural equation (e.g. dimensionality analyses, construct validity) as well and served as a basis for large-scale assessments.

Sample and procedures

A pilot study was conducted with the participation of 84 students. The online colour perception and interpretation test was completed two first (N = 43; girls: 51%) and two classes of second-grade students (N = 41, girls: 53%) with a mean age 8,24 (SD=0,75).

All the online tests presented in the developmental experiment were constructed and carried out through the eDia (Molnár & Csapó, 2013, 2019; Csapó & Molnár, 2019). The realisation of studies was supported by the Center for Research on Learning and Instruction, University of Szeged. The tests completion required a computer, a standard web browser, Internet connection, and headphones for lower primary school students for multiple listening to the test instructions.

Results of a pilot tests

The internal consistency of the test (84 items) proved to be high (Cronbach-α = 0.906). However, after scaling “alpha if item deleted” and taking into account modification and fit indices extracted from analyses with structural equation techniques after deletion of 21 items we achieved improved reliability (Cronbach-α=0,92). These changes resulted in higher reliability on the whole and on subtest level. Considering the test difficulty the mean achievement was near 67% among first and second-grade students. The originally constructed theoretical model and moderate strong correlation coefficients raised a need for dimensionality analysis of the examined construct. The resulting data supported the validity of the constructed theoretical framework. The empirical data revealed significantly better fit on the four-dimensional theoretical model than on the one-dimensional model ($\square^2$=158,74, df=6, p<0,001). On this basis of results, the theoretical model of colour perception and interpretation

are defined by four dimensions: colour sensitivity, colour-and form recognition, colour memory, and colour and meaning.

Significant differences are detected between two age groups (first and second-grade students) in all subtests ($t=–4,7$, $p<0,001$), indicating the differences related to the development of colour perception and interpretation. The empirical findings did not reveal gender-related significant differences. Summarizing the results, the test and sub-tests are reliable, the theoretical four-dimensional model is valid, but due to sample size this finding we could not fully generalize this finding. The correlation analyses revealed a moderate-strong relationship between the subscales. Considering the difficulty level of the test, the items almost covered the construct levels of the examined sample, but future test development should focus on more difficult items.

A comparative pilot study of creation and interpretation skills and abilities related to colours

Little theoretical or empirical work has been conducted to date on correlations between colour perception and creation. In the nineties of the last century in Hungary, the Leonardo Program attempted to understand skills and abilities related to colour perception. The Visual Narrative Test (VNT) applied by this program involved subtests for the assessment of colour perception and interpretation. The task divided colour perception tasks into three subtest levels: (1) content independent, imaginative, decorative, colouring, (2) colouring related to associations on content (without taking into account the compositional possibilities), (3) using colour to depict functions in composition, colour moods are clearly recognizable (Kárpáti, 1996). Interrelations among the three subtests were revealed, but colour perception and creation with colours were found to be different skill sets.

Sample and procedures

The sample of the pilot study consisted of two classes of second-grade students (N=54, boys: 59,3 %, N=32), with a mean age 8,65 (SD=0,48). The online test comprised 71 items. The test battery was expanded with two colour sensitivity and five colour and meaning tasks. Both tests consisted of the following substests: recognition of main colour, observation of colours on works of art, expressive use of colour theory knowledge, observation of colour and light (colour sensitivity). Students completed the test within 45 minutes. Both data collection were carried out with the help of an administrator, in a classroom setting. The online test was carried out through eDia system. The auxiliary materials applied in the preparation of administrators are located in annexes of thedissertation.

Results

Children perceived the tasks enthusiastically and found it challenging both in the online and traditional environments. Newly developed online items did not improve the behaviour of the test. At the same time, the reliability of colour-and form recognition subtest transmitted without change proved to be significantly lower on this second pilot measure. As a possible reason, it may be that compared to students' achievement with average 68%, and during this

pilot assessment students achieved results with an average near to 85% average significantly lower standard deviation (SD=14,8% average points).

Results indicate a weak correlation between perception and creation with colour, which means that it the assumption that those who perform better in perception are equally successful in creation or vice versa. The generalizability of results is also limited by the small sample size, as well as evaluation human resource requiring evaluation, containing subjective elements.

Asessment of colour perception and interpretation in educational environment on fifth-to seveth grade students: test development and related pilot studies

In this part of the thesis, the construction of online assessment tasks of higher elementary students in fifth and sixth-grades are described. Methods and major reports on behaviour of the test will be outlined.

The test was completed with the participation of 51 students (first-grade, N=20, boys: 60%; sixth-grade, N=31, boys: 29%) and consisted of 61 items. Students' mean age was 11,51 years (SD=0,61). Students completed the test within 45 minutes. The assessment took place in schools' ICT rooms, using school infrastructure. The reliability of the whole test proved to be good (Cronbach-α =0,84). The analyses of subscales supported the assumption that on possibility of gender-related differences in higher elementary school students. The test suitable for online assessment because of its relatively low reliability and easy-to solve tasks required further development.

Research progress test development of online assessment of higher elementary students in sixth and seventh-grades, second experiment: methods and results

Completion of pilot assessments among fifth and sixth-grade students was followed by data analyses and test item correction and conducted a repeated measure. In the pilot study participated altogether 94 students (sixth-grade, N=43; boys 44,2 %; seventh-grade, N=51; boys 58,8 %), with a mean age 12,39 (SD=0,58). Students completed the test and a background questionnaire (14 items) within 45 minutes. The pilot study was completed in November 2017.

The questions concerned pupils' attitudes towards tasks, their computer use habits, their school grades and their basic knowledge related to colour theory. Like in previous assessments, the four dimensions of colour perception and interpretation were measured: colour sensitivity (37 items), colour-and shape recognition (19 items), colour memory (9 items) and colour and meaning (10 items). The students solved the tasks in schools' ICT rooms. The online tasks were administered and carried out through the eDia system. After revision, the reliability of the test improved significantly (Cronbach-α=0,86), but the item difficulty seems to be still weak.

Despite our endeavours in test difficulty, the test still proved to be easy to solve, the average achievement of students reached near 80%, strongly differentiated on subtests level. Behind the ceiling effect should be the fact that correcting the difficulty of the items is generally a challenging task. After we constructed the items for subtest basis on curriculum analysis, we avoid emphasis on disciplinary knowledge. As a result of the analyses -, similarly to those found in lower school grade sample, we obtained moderate correlation coefficients (r = 0.33-0.65),

which confirmed the assumption that the test measures related but not completely identical skills and abilities.

The originally constructed theoretical model and moderate strong correlation coefficients raised a need again for dimensionality analysis of the examined construct. The resulting data supported the validity of the constructed theoretical framework on higher grade sample, too. The empirical data revealed significantly better fit on the four-dimensional theoretical model than on the one-dimensional model (χ^2=26,25, df=6, p<0,001).

According to model fit indices and findings of first and second pilot measurements we confirmed the assumption that the construction of colour perception and interpretation is defined by four subskills, presented in previous chapters. Based on these findings, H1 and H2 hypotheses are optionally confirmed. Whereas we applied the methods of item response theory, we concluded that based on values of the person/item map from the aspect of future test development, it is recommended to add more difficult tasks in all four dimensions of the construct being tested. From all examined background variables the strongest correlation were detected between attitude towards tasks and levels of interpretation of colours (r=0,38; p<0,01), image editing software or application habits and levels of interpretation of colours (r=0,33; p<0,01), and between participation on drawing-or other contest and image editing software or application habits (r=0,30; p<0,01). The performance of students who admit colour deficiency (3,2% of the sample) did not significantly differ from others. According to the results of two-sample t-tests, differences were significantly detectable among boys and girls, in favour of achievement of girls. The second pilot results confirmed the assumption about the possibility of generating a reliable, valid instrument, feasible for assessment of colour perception and interpretation.

The results of the pilot measurements raised new questions regarding the development of the construct and the structure's possible changing nature (H3). Nevertheless, the results are not consistent with the performances of the individual subscales of the sample (e.g. differences between boys and girls) – H4 and H5 hypotheses.

In order to maintain more precisely features in understanding the structure of the construct, aiming to detect changes in development large-scale assessments were conducted. Further results will be discussed in the next chapters of the dissertation.

LARGE-SCALE ASSESSMENT OF STRUCTURE AND DEVELOPMENT OF COLOUR PERCEPTION AND INTERPRETATION AMONG SECOND AND SEVENTH GRADE STUDENTS

We developed an online test battery that may provide information about the development of colour perception and interpretation in a wide age range. Pilot data collected on small samples served as a basis for the development of research questions related to the properties of the construct such as dimensionality or structure of the assessed construct. However, large-scale assessments are needed to refine the tests and report on developmental changes on the assessed

construct. In this chapter of the dissertation, the results on large-scale assessments executed among two and seventh-grade students will be discussed.

Methods at large-scale assessments

Sample

7087 second and seventh-grade students participated in this study. The first data collection involved second-grade students from 142 Hungarian schools (N=4183, boys: 50,1 %) tested between January to March 2017. The second assessment in the large sample was conducted one year later, with the involvement of seventh-grade students from 99 Hungarian schools (N=2904, boys: 49,1 %). Both large samples belong to a longitudinal research program of MTA-SZTE Research Group on the Development of Competencies, a part of the large samples who registered for testing and were able to provide all the conditions for an online test completion (headsets for multiple listening to the instructions, ICT classroom, internet connection).

Design of large-scale assessments, procedures

In both cases, the tasks were constructed and administered through the eDia (Electronic Diagnostic Assessment) system. While second-grade students had 45 minutes for completion, seventh-grade students had only 20 minutes to finish their tasks. Students completed the tests in schools' ICT rooms, using school infrastructure. The test instructions for lower primary grade students were narrated, they could click on the headset icon and listen to it multiple times. Taking into account the computer skills such as keyboard and mouse usage skills of students in lower primary education, they could click and drag the items (Molnár & Pásztor, 2015). The issue of stabilization of colours was solved with *Web Safe Colour Palette*.

The empirical data were analysed both with classical test theory (SPSS software), item response theory (ConQuest software Bond & Fox, 2013; Molnár, 2013). Techniques for building structural equation were applied by using MPlus software.

The Cronbach-α was used for the primary testing of the reliability of the assessment instrument.

Difficulty levels of the test items and students were compared and visualized with person/item maps. In order to reveal the structure of the construct, we ran Confirmatory Factor Analysis.

For the dichotomus data fit on the theoretical model we applied WLSM (Weighted least squares mean and variance adjusted) and THETA parametrization (Muthén & Muthén, 2010). Adjudging model fit was framed by $\chi 2$ statistics mentioned above. CFI (Comparative Fit Index), TLI (Tucker-Lewis Index) and RMSEA (root mean square error approximation) fit indices were applied. While in cases of CFI and TLI indices cut-off value were desirable above 0,90, recommended by Bentler (1990), RSMEA cut-off values should be below 0,08 (Browne & Cudeck, 1993; Fan & Sivo, 2005; Vandenberg & Lance, 2000).

With regard to the dimensionality analyses rest on comparison of nested models, we could not apply traditional χ^2 test statistics. Instead, we applied a special DIFFTEST command and χ^2 test implemented in MPlus program (Muthén & Muthén, 2010).

In order to compare the results of second and seventh-grade students, firstly we examined the test performances with anchor items. By applying item response theory, the students were placed on a common logit scale. Differences related to school grade and gender were estimated with two-sample t-test, background variables and achievement on colour perception and interpretation test was compared with correlation analysis.

Basic statistical parameters among grades

Supporting the results of pilot studies and based on basic statistical parameters estimated and the dimensional person/item map we concluded that both tests developed for second and seventh-grade students were not difficult to solve. Nevertheless, the reliability values proved to be high, therefore the students' construct levels were similarly characterized.

Correlations among subscales of achievement and structure of the assessed construction on second and seventh-grades

Correlations among subscales of achievement and structure exhibited low correlation coefficients (r=0,12-0,34), therefore we assumed, that the subscales adequately distinguish between subskills. Construct validity was tested using a system of structural equation splitting the sample into two school grades. Based on the evaluation model fit, it seems obvious that the construct of colour perception and interpretation is not unified, it could be defined by four subskills: colour sensitivity, colour-and form recognition, colour memory and colour and meaning. After confirming the hypothesis related to the four-dimensionality of the construct has been confirmed (χ^2=1275,7, df=6, p<0,001), we ran a four-dimensional person/item analysis. By separating dimension into distribution curves the analysis revealed that the Colour and meaning subtest differed the most the second-grade students. It was followed by Colour-and form recognition. The weakest deviations were detected in Colour sensitivity and Colour memory subtests. Identical analyses were performed on the seventh-grade tests and they revealed moderate correlations among the performances on subtests.

The correlation coefficients could be defined as moderate-strong, indicating that the substructures of the construct parallelly changes over time progress, increasingly influencing each other. The strong-moderate correlations also indicate that despite the reliability of the test which indicated that we assessed one construct, the subtests did not measure the same, unified construct. The performances on subtests are not interchangeable. Therefore, we performed a construct validity analysis. The resulting data confirmed the theoretical model, that the construct of colour perception and interpretation could be defined by four subskills. Fit indices of the one-dimensional model were under the required values (CFI=0,854, TLI=0,850, RMSEA=0,031), but within the modification indices became acceptable.

The validity of the fourth dimensional model has been revealed (χ^2=363,0, df=6, p<0,001). Afterwards, we followed the same procedures as previously, we ran a four-dimensional person/item analysis to explore the distribution of the construct among seventh-

grade students. The four-dimensional person/item map did not explore spectacular differences in performances of seventh-grade students, but the Colour and meaning subtest is still disposes of the most significantly differential property in both grades.

Predicting power of some subtests in relation to the development level of colour perception and interpretation among grades

In accordance with the four-dimensional model, we built a structural equation model and performed the examination of the predicting power of the construct of colour perception, now distributed into second and seventh-grade students' data. Results fit significantly better on the four-dimensional theoretical model than the one-dimensional model. This result had been confirmed for both grades. Correlations among development levels in second grade were low, but in seventh grade they were moderate to strong. These results are supported by pathways regressions in SEM model. These findings supported the assumption that the construct is changing over time. To support this assumption, we performed invariance analysis using the anchor items.

Invariance vs. variance of the colour perception and interpretation construct with time progress

In the basic model, these analyses were performed only with the 29 anchor items, starting from the empirical model of the four-dimensional construct. Adjudging basic model fit (CFI=0,929, TLI=0,922, RMSEA=0,055) almost adjusted criteria of well model fit, within integration parts of the modification index suggestions into the model, it had been improved significantly better model fit (CFI=0,966, TLI=0,963, RMSEA=0,038), therefore we performed exploratory confirmatory factor analysis where we correlated these three items. Whilst we worked with categorical variables, therefore the measurement invariance was tested in several steps, by exploratory factor analyses. In spite of the absence of invariance among second and seventh grades, which requires further analysis, we assumed the existence of invariance among second and seventh grades, therefore the comparability of the performances among these two grades on a common logit scale.

Development the structure of colour perception and interpretation among second and seventh-grades: placing test performances on a common logit scale

After scaling "alpha if item deleted" option in SPSS, left 103 items were placed on a common logit scale, therefore Rasch-analyses were performed among second and seventh-grade students' test performances. From 109 items 19 were administered among second-grade students, 29 among second and seventh, and 55 items were administered on seventh-grade school students. The reliability coefficients, EAP reliability=0,812, therefore it had been evidenced that the test bank is homogenous and adequate for assessment of the construct of colour perception and interpretation. Based on the person/item map we could conclude that the anchor items worked well. They include easy, very easy and average tasks, therefore the assumption about the difficulty is also supported. The items that dispose of a low number are easier to solve than the items which dispose of higher numbers. After placing the test

performances on common logit scale, we applied two-sample t-test among the second (M=1,93, SD=1,01) and seventh-grades (M=2,33, SD=1,01) students expressed in logits (F=3,88, p=0,49, t=-16,39, p<0,001). These findings supported the hypothesis that there are significant differences among the test achievements among second and seventh-grade students in favour of seventh-grade students.

The distribution curves presented on the common logit scale among second and seventh-grades revealed that both grades performed at almost the same level. This means that students of lower and higher levels performed on almost the same logit scale, the two logits scales almost covered the same one for both. The difference is detectable in the proportion of students at each logit level. The relevance of previously validated four constructed theoretical model we ran four-dimensional of test performance and placed on a common logit scale. As was the case in previous analyses, the students deviate similarly in four dimensions, therefore we could conclude that the items still seem easy-to-solve.

Gender-related identities and differences in colour perception and interpretation development

First, analyses in gender distribution were performed on test and subtest level, in a traditional way, in percentages. In order to maintain the differences among grades, we worked with scaled data. Girls outperformed boys on the whole and subtest level. Significant differences were revealed among second and seventh grade students in favour of girls.

Correlation of development of colour perception and interpretation and background variables

In general, there is a weak to moderate correlation between both grades related to the development of colour perception and school marks. We revealed no background variable that would be strongly related, our would have more significant determinant power on students' colour conception and colour interpretation development. However, the time spent on completing tasks among seventh grade students influenced achievements. Those who spent more time on solving the tasks generally performed better on the test (the correlation is still weak, r=0,29, p<0,001).

Analysis of possible influencing factors on the development of colour perception and interpretation among second and seventh-gade students

Correlation between the variables of test performance and background variables were discussed in the previous chapter. The examined variables are related to and convey the effects in different degrees. As a result, if only the correlations were examined, the influencing effects would not be eliminated, but they continue to appear in the correlation coefficients. However, regression analysis with partial correlations would exclude multiple relationships within the context. According to the results of regression analysis, it revealed that there was a significant effect on gender and mark from mathematics among second-grade students, explaining the 15% of the variance. In seventh-grade, 34% of the variance was explained after the non-significant effects were eliminated. Among the background variables, gender (5%) and colour and meaning interpretation (6%), and the time spent on completing tasks (5%) proved to be important.

GENERAL DISCUSSION

Research objectives of the dissertation were realised. Online assessment colour perception and interpretation among elementary school children can be realised in school settings. The first research objective was achieved in the literature review. The second research objective was fulfilled after demonstrating the feasibility of developing a test battery for online assessment of colour perception and colour interpretation, suitable in educational context, providing immediate feedback to teachers and has an objective, reliable psychometric properties.

The resulting data of large-scale assessment confirmed the hypothesis (H1) that developing a reliable test battery for an assessment of colour perception and interpretation in educational environment, among wide age ranges is possible, providing personalized feedback to teachers. Although the elaborated test system requires further revision (item difficulty/developing further with more difficult items on all four examined subtests), overall, the elaborated tests are reliable both on test and subtest level. The online test constructed for online assessment is feasible for early stages of school and provides teachers a rapid feedback about their students' performance in key areas of competence in workplace and private life.

The empirical data supported the hypothesis about the four-dimensional feature of the construct of colour perception and interpretation, which has been confirmed by applying models of structural equation. Hence, the H2 hypothesis is also confirmed, namely, that the colour perception and interpretation, as a component of the system of visual skills, could be defined by four dimensions. The construct consists of colour sensitivity, colour-and form recognition, colour memory and colour and meaning. These four components are strongly related. The invariance analyses highlighted that although the four-dimensional model is valid on both grades, the subscales are separated, play a different role and their developmental level is changing over time. Whereas in second grade, colour sensitivity and colour-and form recognition subscales have been more dominant, on seventh-grade the all four components play almost the identical role in determining the whole structure of the developmental level.

Analyses of the development of the construct have confirmed the assumption that colour perception and interpretation significantly develop among second and seventh-grades (H3). This hypothesis is also supported and consistent with other results reported on development of colour and meaning, assessed through creation (Gaul-Ács & Kárpáti, 2018; Kárpáti, 1996). The areas of colour sensitivity (Sugita, 2004), colour-and form recognition (Gathers et al., 2004; McGivern et al., 1997), colour memory (Mecklenbäuker et al., 2001) and colour and meaning (Kárpáti, 1996) are proved to develop significantly over years. Reflecting on development, Arnheim's (1986) hypothesis is confirmed, namely that picture always dominates over the cognitive aspect of experiences, which is consistent with Sipe's theory that perception is a sensorial equivalent of cognitive level (Sipe, 2008; Willson, Falcon & Martinez, 2014).

Limitations of the analyses are ceiling effect caused by easy tasks. Our gender-related results revealed consistency with colour sensitivity subetsts of previous research efforts (Gaul-Ács & Kárpáti, 2018; Kárpáti, 1996. Regarding the colour sensitivity subtests, girls performed better both on second and seventh-grades, which is consistent with findings of Abramov et al. (2012). Regarding girls' results, they performed better in seventh-grade, but it was not the case

with the boys' performance. There were no significant differences among boys in second and seventh grades (H3). At the same time, ANOVA analyses showed no significant differences among girls in second grade and boys in seventh grade (H4). In general, considering the findings, the present research is well-grounded for further examination of visual skills, especially in detecting gender-related differences.

Answering the question of the sixth research objective, we did not find a background variable that would significantly determine the development of colour perception and interpretation. On the second-grade, we found some correlation between school marks in mathematics and gender, as it was identified on seventh-grade, but the impact of these coefficients was negligible. Overall, we could explain the 15% and 34% of the variance. In this case, this hypothesis (H5) cannot be fully rejected.

Results indicated that the test had week differential power in high skill levels, which constraitly limits the generalizability of results. However, after revision, the test can be extended for assessment among preschool children. Taking into account the lessons of tests administered in higher grades, the interpretation of colour among higher school grades should be assessed with creative tasks.

The dissertation justified the statement about the possibility of online assessment of colour perception and interpretation in lower school grades, providing rapid and objective feedback to teachers about their students' performances in this area.

Our research provides a framework of colour perception and interpretation skills relevant for teaching Visual culture, the Hungarian discipline for art and design education. It mapped the structure of the colour perception and interpretation construct, through which we could gain a data about elementary school children's perception and interpretation of colours.

PEDAGOGICAL RELEVANCE OF THE RESARCH

Analyses performed in the dissertation can serve as a basis for developing a more effective educational program related to teaching about colour in arts education curriculum with more precisely defined content to be taught, areas for development and achievement targets for school age-groups. This instrument provides immediate and personalized feedback for art teachers about results of their educational efforts in this area much more precisely than the observation of artworks of pupils.

Our research may contribute to a new model of child art development, too. The research design follows the methods of curricula based research of visual arts abilities and skills for a competency model. The hypothetical model consists of four subskills derived from the Hungarian curriculum and from the scientific literature, explored in the theoretical background. Besides other visual skill tests, our instrument also belongs to the assessment repertoire of a longitudinal curriculum design project, the Moholy-Nagy Visual Modules (2016-2020). This project aims to integrate visual communication, media culture, environment culture and contemporary art in an innovative, modular art education curriculum and use innovative, computer-based assessment.

REFERENCES

Abramov, I., Gordon, J., Feldman, O., & Chavarga, A. (2012). Sex and vision II: color appearance of monochromatic lights. *Biology of sex differences*, *3*(1), 21.

Albers, P., & Harste, J. C. (2007). The arts, new literacies, és multimodality. *English Education*, *40*(1), 6-20.

Arnheim, R. (1986). *Art and visual perception: A psychology of the creative eye*. Berkeley: University of California Press.

Babály, B., Budai, L., & Kárpáti Andrea (2013). A térszemlélet fejlődésének vizsgálata statikus és mozgó ábrás tesztekkel. *Iskolakultúra, 13*(11), 6-19. Retrieved from http://epa.oszk.hu/00000/00011/00179/pdf/EPA00011_iskolakultura_2013_11.pdf

Bakos T, Bálványos H, Preisinger, Zs. & Sándor Zs. (2000). *A vizuális nevelés pedagógiája.* Budapest: Balassi Kiadó.

Bentler, P. M. (1990). Comparative fit indexes in structural models. Psychological Bulletin, 107(2), 238-246. Retrieved from http://dx.doi.org/10.1037/0033-2909.107.2.238

Bintz, C. (2016). Visual literacy: does it enhance leadership abilities required for the twenty-first century?. *Journal of Visual Literacy*, *35*(2), 91-103.

Bond, T. G. & Fox, C. M. (2013). Applying the Rasch model: Fundamental measurement in the human sciences. Psychology Press.

Burkitt, E. (2004). Drawing conclusions from children's art. *PSYCHOLOGIST-LEICESTER-*, *17*, 566-569.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage focus editions*, 154, 136-136.

Cowan, N. (Ed.). (1997). *The development of memory in childhood.* Hove, UK: Psychology Press

Cox, M. (2005). *The pictorial world of the child*. Cambridge University Press.

Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia System. *Frontiers in psychology*, 10, 1522.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual review of psychology*, *65*, 95–120.

Gathers, A. D., Bhatt, R., Corbly, C. R., Farley, A. B., & Joseph, J. E. (2004). Developmental shifts in cortical loci for face and object recognition. *Neuroreport*, *15*(10), 1549.

Gaul-Ács, Á & Kárpáti, A. (2018). *Óvodás gyermekrajzok vizsgálata a három narratív rajz képalkotó feladatsorral. Magyar Pedagógia, 118(3). 279-306.* DOI: 10.17670/MPed.2018.3.279

Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology*, *10*(13), 805-808.

Kail, R. (1990). A series of books in psychology. The development of memory in children, 3rd ed. New York: W H Freeman/Times Books/ Henry Holt & Co.

Kárpáti, A. (1996). A Leonardo Program hatásvizsgálata. Magyar Pedagógia, *96*(1), 3-34.

Kárpáti, A. (2011). Esztétikai nevelés az új képkorszakban: a gyermekrajztól a vizuális nyelvig. *Magyar Tudomány, 9*, 1058-1064.

Kárpáti, A., Babály, B. & Simon, T. (2015). Az eDia online tesztrendszer pilot kísérletei a Térszemlélet és Vizuális kommunikáció területén. In B. Csapó, & A. Zsolnai (Eds.), *Online diagnosztikus mérések az iskola kezdő szakaszában*. (pp. 29-58). Budapest: Oktatáskutató és Fejlesztő intézet.

Kárpáti, A., & Gaul, E. (2011). A vizuális képességrendszer: tartalom, fejlődés, értékelés. In B. Csapó és A. Zsolnai (Eds.), *Kognitív és affektív fejlődési folyamatok diagnosztikus értékelésének lehetőségei az iskola kezdő szakaszában* (pp. 41−82). Budapest: Nemzeti Tankönyvkiadó.

Kárpáti, A., & Pataky, G. (2016). A Közös Európai Vizuális Műveltség Referenciakeret. *Neveléstudomány*, (1), 6-21. Retrieved from http://nevelestudomany.elte.hu/index.php/2016/04/a-kozos-europai-vizualis-muveltseg-referenciakeret/

Kárpáti A., & Pethő, V. (2012). A vizuális és zenei nevelés eredményeinek vizsgálata. In: Csapó, B. (Ed.). *Mérlegen a magyar iskola.* (pp. 451-483). Budapest: Nemzeti Tankönyvkiadó.

McGivern, R. F., Huston, J. P., Byrd, D., King, T., Siegle, G. J., & Reilly, J. (1997). Sex differences in visual recognition memory: support for a sex-related difference in attention in adults and children. *Brain and cognition*, *34*(3), 323-336.

Mecklenbräuker, S., Hupbach, A., & Wippich, W. (2001). What colour is the car? Implicit memory for colour information in children. *The Quarterly Journal of Experimental Psychology: Section A, 54*(4), 1069-1086. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/713756006?needAccess=true

Ministry of Human Resources. (2012a). *Kerettanterv – Vizuális kultúra (1–2. és 3–4. évfolyam).* Budapest: Ministry of Human Resources. Retrieved from http://kerettanterv.ofi.hu/01_melleklet_1-4/index_alt_isk_also.html

Ministry of Human Resources. (2012b). *Nemzeti alaptanterv – Vizuális kultúra (1–12. évfolyam).* Magyar Közlöny, (66), 10798–10812. Retrieved from https://ofi.hu/sites/default/files/attachments/mk_nat_20121.pdf

Molnár, G. (2002). A tudástranszfer. *Iskolakultúra*, *12*(2), 65-74.

Molnár, G. (2013). *A Rasch-modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában.* Budapest: Gondolat Kiadó.

Molnár G., & Csapó, B. (2013, April). Az eDia online diagnosztikus mérési rendszer. Paper presented at XI. Pedagógiai Értékelési Konferencia. Szeged. Abstract retrieved from http://www.edu.u-szeged.hu/pek2013/download/PEK2013_kotet.pdf

Molnár, G. & Csapó, B. (2019). Technology-based diagnostic assessments for identifying early learning difficulties in mathematics. In Fritz-Stratmann, A., Räsänen, P. and Haase, V. (Eds.), *International Handbook of Mathematical Learning Difficulties* (p.683-707). Heidelberg: Springer.

Molnár, G., & Pásztor, A. (2015). A számítógép alapú mérések megvalósíthatósága kisiskolás diákok körében: első évfolyamos diákok egér-és billentyűzet-használati képességeinek fejlettségi szintje. *Magyar pedagógia*, *115*(3), 239-254.

Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide.* Los Angeles, CA: Author.

Pataky, G. (2012). Vizuális képességek fejlődése 6-12 éves korban a tárgykultúra tanításának területén. Budapest: ELTE TÓK.

Schönau, D. W. (2012). Towards developmental self-assessment in the visual arts: Supporting new ways of artistic learning in school. *International Journal of Education through Art*, *8*(1), 49-58.

Sipe, L. R. (2008). *Storytime: Young children's literary understanding in the classroom*. New York and London: Teachers College Press.

Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling*, *12*(3), 343-367.

Siu, K. W. M., Lam, M. S., & Wong, Y. L. (2015). Gender Differences in Children's Use of Colors in Designing Safety Signs. *Procedia Manufacturing*, *3*, 4650-4657.

Sugita, Y. (2004). Experience in early infancy is indispensable for color perception. *Current Biology*, *14*(14), 1267-1271.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, *3*(1), 4-70.

Vurro, Ling, & Hurlbert (2013). Memory color of natural familiar objects: Effects of surface texture and 3-D shape. *Journal of Vision*, *13*(7), 1-20.

Wagner, E., & Schönau, D. (Ed.). (2016*). Cadre Européen Commun de Référence pour la Visual Literacy-Prototype Common European Framework of Reference for Visual Literacy-Prototype Gemeinsamer Europäischer Referenzrahmen für Visual Literacy-Prototyp.* Münster: Waxmann Verlag.

Willson, A. M., Falcon, L. A., & Martinez, M. (2014). Second graders' interpretation of character in picture Book Illustrations. *Reading Horizons (Online)*, *53*(2), 1-26.

Zentner, M. R. (2001). Preferences for colours and colour-emotion combinations in early childhood. *Developmental Science*, *4*(4), 389–398. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/1467-7687.00180/full

PUBLICATIONS RELATED THE DISSERTATION

Gaul, A., Kárpáti, A. & Tóth, A. (in press). Visual literacy development of young children: nurture above nature. *International Journal of Education through Art.*

Pásztor A., Babály B., Simon T., & Tóth A. (2016, June). *A kombinatív és a vizuális képességek összefüggései 5. osztályban – egy pilot vizsgálat eredményei*. Paper presented at the 1. Művészetpedagógiai Konferencia, Budapest. Abstract retrieved from http://mpk.elte.hu/download/MPK_2017_magyar_kotet_final.pdf

Tóth, A. (2015, November). *A színek észlelése, mérési lehetőségek*. Poster session presented at the XV. Országos Neveléstudományi Konferencia, Budapest. Abstract retrieved from http://onk2015.conf.uni-obuda.hu/wp-content/uploads/2015/01/ONK_2015_tartalmi_osszefoglalok.pdf

Tóth, A. (2016a, April). *A literature review and curriculum analysis for color aesthetic experience development assessment among 6- to 12- year-old children*. Paper presented at the 14th Conference on Educational Assessment, Szeged. Abstract retrieved from http://www.edu.u-szeged.hu/pek2016/wp-content/uploads/2015/11/PEK2016_kotet_v.pdf

Tóth, A (2016b, July). *Framework of colour perception and appreciation: aesthetic and non-aesthetic experiences of colour.* Paper presented at the EARLI JURE Conference, Helsinki, Finland.

Tóth, A, & Kárpáti, A. (2016a, September). *Assessment of colourperception and interpretation*. Paper presented at the InSEA Regional Conference, Vienna, Austria. Abstract retrieved from https://www.dieangewandte.at/jart/prj3/angewandte-2016/data/uploads/Aktuelles/Veranstaltungen/2016/InSEA_Booklet_2016.pdf

Tóth, A, & Kárpáti, A. (2016b, October). *Colour perception and interpretation framework.* Paper presented at the 48th Annual International Visual Literacy Conference, Montreal.

Tóth, A. (2016c, November). *Színpercepció és színalkotás: a vizuális alkotói és befogadói képességrendszer összehasonlító vizsgálata.* Paper presented at the XVI. Országos Neveléstudományi Konferencia, Szeged. Abstract retrieved from http://www.edu.u-szeged.hu/onk2016/download/ONK_2016_program_es_absztraktkotet.pdf

Tóth, A (2017a). Az Én Ujságom: gyermekkép, képi ábrázolás, pedagógiai üzenetek, *Új Kép folyóirat*. *19*(1-4), 29-36.

Tóth, A (2017b). A színpercepció és színértelmezés mérésnek tartalmi keretei általános iskolás diákok körében, *Iskolakultúra, 27*(1-12). 34-47. DOI: 10.17543/ISKKULT.2017.112.34

Tóth, A (2018, April). *A színpercepció és színértelmezés teszten elért eredmények és a háttérváltozók közötti összefüggések vizsgálata 6. és 7. évfolyamon*. Paper presented at the 16th Conference on Educational Assessment, Szeged. Abstract retrieved from http://publicatio.bibl.u-szeged.hu/13740/1/PEK_2018_PROGRAMME_ABSTRACTS.pdf

Tóth, A (2019, June). *Teaching colours: curricula analysis and framework for assessment of colour perception in educational context.* Paper presented at 3rd ELTE Workshop for Arts Education – EAWE, Budapest. Abstract retrieved from http://mpk.elte.hu/en/download/EWAE_2019_Proceedings.pdf

Tóth, A, Kárpáti, A & Molnár, G. (2017). A színpercepció és színértelmezés online mérésének lehetőségei kisiskolás korban. *Magyar Pedagógi*a, *117*(4), 399–421. DOI: 10.17670/MPed.2017.4.399

Tóth, A, Kárpáti, & Molnár, G. (2019). A színpercepció és színértelmezés mérése 7-9 éves diákok körében. *Iskolakultúra*, *29*(1), 17-28. DOI: 10.17543/ISKKULT.2019.1.17

Tóth, A, Kárpáti, A & Molnár, G. (submitted). Teaching about colour – the legacy of the Bauhaus masters. *Visual Inquiry, Intellect Books.*

Tóth A., Simon T., & Kárpáti A. (2017, April). *Moholy-Nagy Vizuális Modulok: beszámoló a színpercepció és vizuális kommunikációs tesztek pilot vizsgálatáról 1., 5. és 9. osztályos tanulók körében.* Paper presented at the 15th Conference on Educational Assessment, Szeged. Abstract retrieved from http://www.edu.u-szeged.hu/cea2017/download/CEA_2017_BOOK_of_Abstracts.pdf

Tóth A., Simon., & Pásztor, A. (2017, June). *Az online színpercepció és vizuális kommunikáció pilot teszteken elért teljesítmények és az iskolai osztályzatok közötti összefüggések vizsgálata*. Paper presented at the 1. Művészetpedagógiai Konferencia, Budapest. Abstract retrieved from http://mpk.elte.hu/download/MPK_2017_magyar_kotet_final.pdf