University of Szeged
Doctoral School of Education

ZSÓFIA GÁL-SZABÓ

SOLVING ENUMERATIVE COMBINATORIAL PROBLEMS IN PRIMARY SCHOOL: THE ASSESSMENT OF PROBLEM COMPREHENSION AND STRATEGY USE

Summary of the PhD dissertation

Supervisor: Erzsébet Korom

Szeged
2020

TOPIC OF THE DISSERTATION

The PhD thesis is concerned with combinatorial reasoning, by which – in line with theoretical approaches established in Hungary (Csapó, 1988; Nagy, 2004) – we mean a theoretical construct consisting of specific operations constituting one component skill of reasoning ability. We use Adey & Csapó's (2012) definition of combinatorial reasoning, which states that combinatorial reasoning is the skill allowing an individual to arrange a set of items according to some pre-determined criteria. As it is defined for our purposes, combinatorial reasoning is analogous with one of the categories of combinatorial problems: enumerative combinatorial problems (see Batanero, Godino & Navarro-Pelayo, 1997). Since the international literature uses the term combinatorial reasoning in a variety of senses and contexts, we have decided to avoid ambiguity by using the term enumerative combinatorial problem in the title and body of this dissertation. The decision is justified by the fact that in studies closely related to our work (see e.g., English, 1991, 1993; Mwamwenda, 1999; Poddiakov, 2011), problems of that type are used for data collection regardless of the theoretical framework in which the study is embedded.

Since Piaget's pioneering work proposing that combinatorial reasoning plays a central role in the development of formal thinking (see e.g., Inhelder & Piaget, 1967; Piaget, 1970), several theoretical and empirical studies have been published on the subject. A number of these (e.g., Mashiach-Eizenberg & Zaslavsky, 2004; Melusova & Vidermanova, 2015; Szitányi & Csíkos, 2015) note the difficulty and challenges of research in this area, in the light of which it came as no surprise when Lockwood (2015) pointed out that we still have a lot to learn about students' combinatorial reasoning and problem solving processes. In connection with solving combinatorial problems, Hadar and Hadass (1981) identified a set of potential errors, of which the appropriate interpretation of the task and the lack of a systematic enumeration strategy are of special interest to us. The aim of the dissertation is therefore to explore learners' interpretation of combinatorial problems and their strategies of enumeration in an effort to contribute to our body of knowledge concerning learners' combinatorial reasoning and to further our understanding of their strategy use.

The introductory section of the dissertation is followed by three chapters discussing theoretical issues, five chapters presenting the empirical studies and finally, a chapter with our conclusions. The dissertation is based on the text of three previously published papers (Gál-Szabó, 2019; Gál-Szabó & Korom, 2018; Gál-Szabó, Korom & Steklács, 2019) and a conference presentation (Szabó & Korom, 2017). The research presented in this dissertation was supported by the Szeged Centre for Research on Learning and Instruction, the MTA-SZTE Science Education Research Group and the Doctoral School of Education at the University of Szeged.

THEORETICAL FRAMEWORK

The majority of studies on combinatorics and combinatorial reasoning place the subject in the context of mathematical thinking and mathematics education (e.g., Csapó, Csíkos & Molnár, 2015; DeTemple & Webb, 2014; English, 2005, 2016; Lockwood, 2013). This approach is supplemented by a general educational or psychological perspective viewing combinatorial reasoning as a component of reasoning ability (e.g., Csapó, 1988, 2001; Inhelder & Piaget, 1967; Nagy, 2004; Zentai, Hajduné Holló & Józsa, 2018). The research programme discussed in the dissertation belongs to the latter of the two approaches. Although the two perspectives are reflected in conceptual differences, the theoretical conclusions and empirical results of one paradigm remain relevant and useful to the other in many cases. As was noted before, combinatorial reasoning is for our purposes (see Adey & Csapó 2012) represented by enumerative combinatorial problems (see Batanero et al., 1997), which means that our research may also have its place among studies concerned with mathematics education.

One obstacle to the comparability of the results of empirical studies on the subject is the circumstance that different test items are used for the assessments. The identification of the type of test item may be assisted by the taxonomy described by Batanero et al. (1997) distinguishing four categories of combinatorial problems: existence problems, counting problems, optimization problems and enumeration problems. For the purposes of our research, studies investigating test items belonging to the category of enumerative problems are of special significance and from the point of view of learners' use of strategies, the results of studies on counting problems are also pertinent. International empirical studies concerned with the above two areas (e.g., Höveler, 2018; Shin & Steffe, 2009) usually evaluate the solution of one or more items involving a single combinatorial operation. In connection with the study of combinatorial reasoning as a system, we know of a number of Hungarian (e.g., Csapó, 2001, Csapó & Pásztor, 2015, Nagy, 2004) studies and studies in some way connected to Hungary (e.g., Wu & Molnár, 2018). Of these, the papers discussed in the dissertation rely on Csapó's theoretical model (Csapó, 1988), which identifies eight combinatorial operations modelling combinatorial reasoning ability: the Cartesian product of sets, variations with repetition, variations without repetition, all variations with repetition, combination with repetition, all subsets, combination without repetition, permutation with repetition.

When evaluating enumerative combinatorial problems, we can analyse – among other things – the end result, i.e., student achievement, and the process of solving the problem, i.e., the method, structure of strategy used by the problem solver to enumerate all arrangements (after Csapó, 2003). Previous Hungarian studies (see e.g., Csapó, 1988, 2001, 2003; Csapó & Pásztor, 2015; Hajduné Holló, 2004; Nagy, 2004; Pap-Szigeti, 2009; Szabó, Korom & Pásztor, 2015) focused on the first of the two aspects of solutions. To our knowledge, there are only international studies addressing the question of the method of enumerating solutions (see e.g., Inhelder & Piaget, 1967; Piaget, 1970; English, 1991; 1993; Halani, 2012; Scardamalia, 1977). The following two paragraphs summarise the literature concerning the above two dimensions of assessment.

A shared property of items included in measurement instruments of combinatorial problem solving is that all possible unique arrangements of elements of the problem set have to be enumerated as required by the set of criteria specified in the problem. When evaluating responses, a method of establishing the accuracy or inaccuracy of the solution may be sufficient for the simplest test items requiring only a few unique arrangements. In more complex cases, however, the above mentioned dichotomous evaluation method results in a great degree of information loss; a more refined evaluation scale is therefore desirable. In what follows, we shall discuss three solutions to this challenge in evaluation. Studies embedded in the Csapó model (e.g., Csapó, 2001; Csapó & Pásztor, 2015; Szabó, Korom, & Pásztor, 2015) typically evaluate responses using the j-index suggested by Csapó (1988). The j-index shows the number of correct combinations (those meeting the criteria), incorrect combinations (those not meeting the criteria) and surplus combinations (correct but non-unique combinations) relative to the number of all possible combinations. The j-index is a good indicator of achievement but it is of less value when the purpose of evaluation is diagnosis. Nagy (2004) suggests a method of evaluating elementary combinatorial ability along four dimensions. The first dimension distinguishes accurate and inaccurate solutions while the remaining three concern the analysis of errors and provide information about the nature (content and quality) and quantity of errors. In their study based on Nagy's model, Zentai and colleagues (2018) suggest a new method of evaluation for younger learners. This method also takes four factors into consideration giving rise to the following four categories of responses: there is at least one correct combination; all possible combinations are enumerated; there is at least one correct combination and all combinations are unique; there is at least one correct combination and there are no incorrect combinations. Researchers recommend the latter two evaluation methods for the diagnostic assessment of elementary combinatorial ability (relatively simple problems) but it is questionable whether they can provide sufficient detail about responses when more complex problems need to be analysed. We may conclude that further proposals are needed for the diagnostic assessment of solutions to relatively complex enumerative combinatorial problems.

With respect to enumerative combinatorial problems, the logic apparent in the enumeration of possible combinations is termed combinatorial strategy after Scardamalia (1977) and English (1991). According to Piaget's theory of cognitive development, the pre-operational stage, the concrete operational stage and the formal operational stage are characterized by different problem-solving strategies (Inhelder & Piaget, 1967; Piaget, 1970). The problem solving behaviour of children at the preoperational state is characterised by random trial and error; next, during the concrete operational stage, some kind of system appears in children's thinking, which later develops fully at the formal operational stage. In accordance with Piaget's findings, English (1991, 1993) identified six increasingly sophisticated strategies used in solving problems involving the Cartesian product of sets ranging from random selection to fully systematic selection (random selection, trial-and-error procedure, emerging pattern, consistent and complete cyclical pattern, odometer pattern and complete odometer pattern. Problem solvers using the most advanced 'odometer' strategies hold one element constant and systematically select all elements paired with it, repeating this procedure for the whole set. English argues that the complete odometer strategy she considers to be the most advanced method is not the only one that may lead to a perfect solution but that is the one she believes to be the most efficient solution. Halani (2012) and Lockwood (2013) investigated the efficiency of problem solving strategies with respect to other combinatorial operations (permutations without repetition and variation with repetition) and came to the conclusion that more than one efficient strategy may be observed for the studied operations but they all share the property of being algorithmic, i.e., they involve some system of enumeration.

Empirical data on strategy use indicate that learners' strategies gradually increase in sophistication with their age (English 1991, 1993). It has also been observed that learners may not employ the same strategy consistently when solving a series of problems but they may switch between different methods of enumerating combinations (English 1991, 1993). There is also some evidence that less sophisticated strategies are accompanied by lower performance while more sophisticated strategies result in higher performance (Csapó & Pásztor 2015).

The methods, the number and types of test items and the sample sizes used in previous small-scale exploratory studies of strategy use in combinatorial problem solving (e.g., English 1991, 1993; Halani, 2012; Palmér & van Bommel, 2018) may be enhanced with the help of technology-based assessment (e.g., eDia, see Molnár & Csapó, 2019). A large-scale analysis of combinatorial strategies is now possible thanks to an algorithm-based classification method (Gál-Szabó & Bede-Fazekas, 2020). The procedure is suitable for the analysis of solutions to problems involving the Cartesian product of sets; it identifies seven categories of strategies based on two characteristics of the enumeration patterns (odometricality and cyclicality): random, slightly cyclical, slightly odometrical, nearly cyclical, nearly odometrical, completely cyclical and completely odometrical strategy.

THE RESEARCH PROGRAMME

The overall aim of our research programme was to further our knowledge concerning combinatorial reasoning and deepen our understanding of the construct. To achieve this aim, we utilised technology-based assessments of combinatorial reasoning focusing on two aspects of the issue: (1) learners' comprehension of the criteria of arrangement (given combinatorial operations) and (2) the method of enumeration while solving the problem, i.e., the strategies employed. Based on the above, the following research objectives were defined for the programme:

(1) To develop and test an online measurement instrument by revising and expanding Csapo's digitised combinatorial test (Csapó & Pásztor, 2015);
(2) To develop a set of variables suitable for the assessment of learners' comprehension of the criteria of arrangement, and to explore the progression of the values of these variables, their association with learner performance and their predictive power among 4th and 6th graders; and
(3) To develop a method of assessment of the use of strategies in solving Cartesian product problems, to explore learners' use of these strategies and analyse the relationship between strategy use and learner performance among 4th and 6th graders.

Three sets of surveys were conducted in order to fulfil our research objectives and the current work presents four studies based on the data collected through the surveys (see Table 1). While the table displays the surveys and experiments in chronological order, the structure of this summary of our studies is determined by the nature of the study objectives.

Table 1. Studies included in the research programme in chronological order

Study	*Objectives*	*Instrument*	*Grade*	*N*	*Data collection*
Eye tracking study	Pilot study for the study of strategy use	CP problem from Csapó's combinatorial test	3rd	48/30	11/2016
Test development	Development and testing of measurement instrument	Eight combinatorial problems	4th 6th	118 121	06/2017
Problem comprehension	Study of problem comprehension	Eight combinatorial problems	4th 6th	482 482	12/2017–01/2018
Strategy use	Study of strategy use in CP problems	Three pertinent problems from the test			

Note: CP: Cartesian product of sets

OBJECTIVES AND HYPOTHESES OF THE STUDIES

Development of measurement instrument

The objective of the study was (1) to develop a computer-based measurement instrument suitable for the assessment of combinatorial reasoning and (2) to test that instrument among 4th and 6th grade students.

When developing our measurement instrument, our aim was to use the online version of Csapó's combinatorial test (Csapó & Pásztor, 2015) as a starting point and construct a measurement instrument that is varied in terms of the contexts of the test items but uniform in terms of the nature and arrangement of those items and in terms of the structure of the instructions. The instrument was to include a variety of Cartesian product problems with different levels of complexity suitable for studying the use of strategies in problem solving.

When testing the measurement instrument, we formulated the following hypotheses (after Csapó & Pásztor, 2015; Szabó, Korom & Pásztor, 2015 and Szabó & Korom, 2016):

- H_1 The test component of the measurement instrument would function appropriately in both target age groups as indicated by reliability indices (Cronbach α)
- H_2 The test component of the measurement instrument is of moderate difficulty and 6th graders would show significantly better performance than 4th graders.
- H_3 The test items would vary in difficulty and the problems involving the Cartesian product of sets would be among the easier items.

Study of problem comprehension

The objectives of the study were (1) to define the variables suitable for the assessment of the comprehension of the criteria of enumeration in problems (given combinatorial operations) and (2) to measure the values of these variables and analyse their relationship with test performance and their predictive power among 4th and 6th graders.

The development and evaluation of the variables, and the data included in our analyses led to the following hypotheses:

- H_1 It is possible to define variables along which learners' comprehension of the criteria of enumeration in the problem (operation) can be assessed and these variables can be unambiguously operationalised for the problems.
- H_2 The variables thus defined can be used to evaluate learners' responses to the test items included in the online measurement instrument.
- H_3 A correlation would be found between the proportion of responses meeting the given criteria and the difficulty of the problem (as shown by the j-index of performance).
- H_4 The values of the variables would display similar patterns in the two grades; answers meeting the criteria of enumeration would occur in higher proportions in the older group.
- H_5 Answers meeting a higher number of criteria would be accompanied by better item and test performance.
- H_6 The proportions of responses meeting the criteria and the explanatory values of the variables would vary across problems involving different operations.
- H_7 The two age groups would display similar patterns with almost equal variances with respect to the explanatory power of the variables.

Eye tracking study

The aim of the study was a thorough, in-depth exploration of the use of combinatorial strategies in problem solving in preparation for future research in that area. In that connection our objectives were (1) to identify combinatorial strategy use, (2) to explore the patterns of responses (the logic of completing diagrams) and (3) to analyse responses and the efficiency of solutions as revealed by fixations on various interest areas among 3rd graders.

We formulated the following hypotheses based on the literature (Csapó & Pásztor 2015; English, 1991, 1993) and our previous experiences:

- H_1 Participants would be characterised by the use of a variety of strategies.
- H_2 A variety of response patterns would be observed with respect to the completion of the diagrams (the construction of the solutions).
- H_3 More sophisticated strategies would be accompanied by better solutions and the solutions of learners demonstrating better performance would be more likely to be characterised by some sort of consistent system of enumeration.
- H_4 Learners performing well on the tasks but using less sophisticated strategies would be characterised by a more careful rereading of their solutions and a greater number of fixations on the response area.
- H_5 Learners performing well on the tasks and using more sophisticated strategies would be characterised by a less careful rereading of their solutions and a smaller number of fixations on the response area.
- H_6 Learners performing the most poorly on the tasks would be characterised by a less careful rereading of their solutions and the smallest number of fixations on the response area.

Study of strategy use

The objectives of the study were firstly (1) to develop a method of converting the raw data to a format allowing the analysis of strategy use using the method detailed in Gál-Szabó and Bede-Fazekas (2020). Secondly, based on the above categorisation algorithm, we undertook to analyse (2) the use of combinatorial strategies, (3) the relationship between strategy use and performance and (4) changes in strategy use for three combinatorial problems involving the Cartesian product of sets among 4th and 6th graders.

Based on the literature (Csapó & Pásztor, 2015; English, 1991, 1993) and the results of our eye tracking study, we formulated the following hypotheses:

- H_1 Students in the two groups would be characterised by varied strategy use.
- H_2 The older group would use more complex (more efficient) strategies to solve the problems than would the younger group.
- H_3 There would be an association between task performance and strategy use: poorer performance would be more likely to be associated with more rudimentary strategies while better performance would be more likely to be associated with more efficient strategies.
- H_4 The most efficient, completely odometrical strategy would be mostly associated with perfect or almost perfect solutions.

H_5 A greater proportion of perfect solutions would be paired with more efficient strategies but more rudimentary, random strategies may occasionally also lead to error-free solutions.

H_6 All three trends in strategy use would be observed across time in the test, i.e., learners may display a decline, stagnation or an improvement in their strategy use.

METHODS OF THE STUDIES

All three surveys of the research programme used the eDia online assessment and evaluation system (Molnár & Csapó, 2019), which relies on computer-based data collection and allows the automatic evaluation and organisation of data in a database. Data collection took place individually for the eye tracking study and in groups, as a classroom activity for the other two surveys. Students' performance was described using Csapó's (1988) j-index. The analysis of the data was conducted using the tools of classic test theory; statistical analyses were carried out in IBM SPSS 24.

Development of measurement instrument

Sample

The sample for the testing of the measurement instrument was recruited through personally approaching school teachers in contact with the MTA-SZTE Science Education Research Group. As a result of the recruitment process, the measurement instrument was tested with 4th (N=118) and 6th (N=121) grade students attending two primary schools in urban areas. Two and three classes of 4th graders and two plus two classes of 6th graders participated in the data collection.

Measurement instrument

Since for the purposes of the research programme outlined in this summary only the test section of the measurement instrument is of relevance, we shall restrict our discussion to that section (all sections of the measurement instrument are described in the dissertation itself). The test contains eight items and it is a revised and extended version of the digital implementation (Csapó & Pásztor, 2015) of the combinatorial test developed by Csapó (2001). The structure of the six pictorial items in the original measurement instrument (type of operation, set sizes and number of elements to be selected) remained unaltered and their order within the test was also kept. This gave us test items involving the following six combinatorial operations (the position of the item in the test is given in brackets): Cartesian product of sets (1-3), all subsets (4), all variations with repetition (5), variation without repetition (6), variation with repetition (7) and combinations without repetition (8).

The revised version kept the original context of three of the items (Items 3, 4 and 8 in our test) and replaced the context of the remaining three items. Two new items were added at the beginning of the test, which allow the evaluation of Cartesian product operations in less complex problems (with smaller set sizes) than the ones already included in the test. Uniform artwork was added to the eight items and the nature and arrangement of the problems and the structure of the instructions were also standardised.

Study of problem comprehension

Sample

The schools participating in a series of surveys conducted in an OTKA research project (Molnár, 2017) at the Szeged Centre for Research on Learning and Instruction were approached and informed they could volunteer to participate in the study. As a result of the recruitment, 35 schools participated from different regions of Hungary. We are in possession of data from 44 school classes in 4th grade (N=790) and 41 school classes in 6th grade (N=751). The analyses discussed in the dissertation were, however, restricted to a subset of the sample ($N_{4th\ grade}$=482, N_{6th} grade=482; the two groups are of equal size by accident).

Measurement instrument

The measurement instrument constructed as a result of our efforts to develop a measurement instrument was used to collect the data. The test section containing eight combinatorial problems was the same as the one tested in the above project with the exception of a few minor adjustments irrelevant to the aims of the research.

Eye tracking study

In a computerised test, an eye tracking equipment allows the recording on video of the process of solving a problem (what the problem solver is doing and when) and of the movements of the eye. It also produces numerical data on processes related to movements of the eye (total fixation time, number of fixations, number of revisits). Data analysis can thus involve the analysis of data both from the video recordings and from the quantified measurements provided by the equipment.

Sample

Two third-grade school classes from a small-town primary school participated in the study with a total of 48 students. Data analysis, however, could not be carried out on the entire sample for the analysis of both the video recordings and the quantified measurements. While all participants in the sample (N=48) could be included in the former, some participants had to be excluded for technical reasons from the latter (reduced sample, N=30).

Measurement instrument

The study used a single pictorial item from the online version of Csapó's combinatorial test (Csapó & Pásztor, 2015), which allowed us to evaluate the process of solving a problem involving the Cartesian product of sets. The task instructed participants to dress a boy using a set of items of clothing (3 pairs of trousers and 4 t-shirts) in all possible distinct ways.

Study of strategy use

Sample

The analyses were carried out on the results of the study on the comprehension of combinatorial problems. As was described above, data were obtained from 482 students in the 4th grade sample and the same number of students in the 6th grade sample.

Measurement instrument

The analyses were carried out on the data from the solving of the first three problems in the test described above. All three items involved the Cartesian product of sets.

THE RESULTS OF THE STUDIES

Development of measurement instrument

(1) The test section of the measurement instrument we developed functioned appropriately with both groups of students (Cronbach-α=0.79 for 4th grade; Cronbach-α=0.74 for 6th grade), and the omission of any of the eight items would result in a decline in reliability (H_1). The item discrimination coefficients indicate that every item discriminated satisfactorily (0.47–0.79 for 4th grade; 0.46–0.79 for 6th grade; $p<0.01$). We therefore concluded that the test was suitable for further research.

(2) The mean test performance was 70.08% (SD=18.26) among 4th graders and 76.53% (SD=14.30) among 6th graders. The measurement failed to support the first part of our hypothesis (H_2) since the test proved to be overall easier than predicted, but – as expected – the older group performed substantially better than the younger group ($p<0.01$).

(3) Participants mean performance for individual test items ranged between 47.19% and 87.32% among 4th graders and between 57.38% and 94.70% among 6th graders, which is in line with our prediction that the items would vary in level of difficulty for students participating in the study (first part of H_3). In addition, as was expected, the problems involving the Cartesian product of sets were among the easier items for both groups of students (second part of H_3). Looking at the ranks of the items with respect to difficulty, we can see that Items 2 and 3 were the easiest and they were followed by Item 1 together with two other items in 4th grade and one other item in 6th grade.

Study of problem comprehension

(1) Taking the nature of combinatorial operations as our starting point, we defined three variables characterising the problem solver's comprehension of the criteria of enumeration. The variables are: set size showing the number of items in a construction, replacement indicating the incidence of repetition, and orderedness related to the order of selection. All three variables may be defined using universal criteria that allow us to decide whether a solution meets a given criterion or not (H_1).

(2) Based on the universal criteria, we defined the specific criteria for the variables created to describe the items in the measurement instrument used for data collection. We further developed a system of codes for the automatic evaluation of the answers providing the values of the variables for each test item (H_2). For the three problems involving the Cartesian product of sets, only the variable of set size was applied owing to the nature of the problems. For Items 4 (all subsets) and 8 (combinations without repetition), the value of the variable of orderedness could not be computed because of the method of recording the answer. That is, set size was analysed for eight items, replacement for five items and orderedness for three items.

(3) A series of paired t-tests revealed that the first three items were the easiest for 4th graders (75.92, 76.04, 76.83%) with no difference in difficulty between them ($p>0.05$). They were followed by Item 7 (65.15%), Item 6 (59.31%) and Item 5 (51.99%), and finally Item 4 (37.57%) and Item 8 (35.50%) proved to be the most difficult to solve. The percentage of answers meeting the given criteria was higher for the five relatively easy items (typically above 80%) than for the three most difficult items (typically below 75%). Of the latter, Items 4 and 5 showed unusually low values (40–60%). The results therefore revealed that the order of difficulty of the test items and the degree of correspondence between the criteria and the

answers were in line with our expectations (H_3). Also, while the older group observed the criteria with greater frequency (45–95% in 6th grade versus 40–90% in 4th grade), the percentages of answers meeting the criteria for individual items showed similar patterns in the two groups (H_4).

(4) Looking at the eight combinatorial problems individually, students on average performed at 80–85% in terms of their answers meeting all of the (1/2/3) criteria (with the exception of Item 8, where the corresponding percentage was around 40–50%). When answers meeting all but one criteria were considered, student performance was typically at 40–50% (a few of the values were lower or higher than that). Performance significantly increased with an increase in the number of criteria met for all eight items in both groups of students ($p<0.01$). Students whose answers met all of the criteria performed better than the average of their group with each of the eight items ($p<0.01$), while the remaining students' performance was lower than average ($p<0.01$) for each problem with the exception of Item 2 in the 4th grade. Looking at the test as a whole a similar pattern could be observed; namely, performance typically increased with an increase in the likelihood of successfully meeting the problem criteria in both age groups ($p<0.01$). That is, answers meeting more criteria were associated with better performance both at the level of an individual item and at the level of the test as a whole (H_5).

(5) The variables explain 5 to 30% of the variance in item performance for the first three problems, where only the criterion of set size is applicable. For the remaining problems, where at least two factors can be considered, the variables explain approximately 50% of the variance in performance in four cases, and the percentage of variance in performance explained by the variables is the highest for Item 5 with a value of around 70%. The contribution of the different variables to the effect varies (3–36%). We may conclude from these results that students' performance was also affected by factors other than the comprehension of the operations but the role of our variables in performance is undisputable. Also, as was expected, although the explanatory power of the variables varies across the items (H_6), the two age groups display similar patterns (H_7).

Eye tracking study

(1) Taking the combinatorial strategies defined by English (1991), we identified a total of 14 strategies in 6 categories, all but one of which were observed in our participants' solutions. Although the different strategies were used with different probabilities, participants were on the whole characterised by the use of a variety of strategies (support for H_1).

(2) The analyses of the video recordings showed that students arriving at perfect solutions were indeed more likely to employ more consistent strategies but perfect solutions could also emerge with rudimentary strategies (partial support for H_2).

(3) The patterns observed in the process of solving the problems differed across the dimensions analysed (columns or rows, from left or from right, from top or from bottom) and none of the dimensions was characterised by a uniform method that applied to every participant completing the diagrams (support for H_3).

(4) Problem solvers using rudimentary strategies but arriving at perfect solutions looked through their answers more carefully at the end of the process, while problem solvers who used more advanced strategies were less likely to check their answers with special care. This pattern is indicated by the number of fixations on the area of the answers: there was a higher number of fixations on this area among students who achieved better results using less advanced

strategies and a lower number of fixations among participants who achieved similar results using more efficient strategies (support for H_4 and H_5).

(5) Similarly to advanced strategy users, students with the lowest performance also paid less attention to their solutions at the end and fixated less on the area of the answers (support for H_6). The explanation presumably lies in these participants not aiming to find all possible solutions; in order to complete the diagrams "as they pleased," there was no need to go over the answers carefully.

Study of strategy use

(1) The first step in building a database suitable for the analysis of strategy use was to reconstruct participants' problem solving behaviour based on the data recorded in the log files during data collection. The log file data included were those pertinent to strategy use (which elements the participant dragged to where and in what order). The events thus obtained were then converted into a format compatible with the strategy classification algorithm (Gál-Szabó & Bede-Fazekas, 2020) to be used for analysis.

(2) While all seven categories of the strategy classification system employed were observed in both age groups included in the study, most answers (80–90%) were characterised by just three strategies: the most rudimentary random strategy (20–30%), and of the more efficient strategies the nearly odometrical (15–20%) and the completely odometrical (35–40%) strategies (partial support for H_1).

(3) Comparing the strategy use of the two age groups there was no difference between them for the first two items ($p>0.05$), while for Item 3, the older group were more likely to use the more efficient strategies ($p<0.05$) (H_2 is supported by the data for Item 3 only).

(4) In line with our predictions, strategy use correlated with performance for all three items in both age groups ($p<0.01$, first part of H_3). The random strategy was used at all levels of performance, while the probability of the three odometrical strategies increased at levels of performance above 60%, and the most frequent strategy used with a performance level of over 90% was the most complex, completely odometrical strategy (this strategy was only observed at a level of performance over 90%). A comparison of the levels of performance achieved with the four strategies occurring with sufficient frequencies (random, slightly, nearly and completely odometrical) revealed that the most rudimentary strategies were associated with lower levels of performance and the more complex strategies with higher levels of performance ($p<0.01$). An increase in performance was therefore indeed accompanied by a greater probability of the use of more efficient strategies (second part of H_3). The levels of performance achieved by participants using the strategies at the two ends of the scale invariably showed a significant difference ($p<0.01$). Students employing the random strategy were on average characterised by performance levels between 55 and 65%, while participants enumerating solutions using a completely odometrical pattern displayed performance levels between 85 and 95%.

(5) With a few exceptions, the most complex, completely odometrical strategy resulted in perfect or almost perfect solutions (H_4). The most rudimentary, random strategy, in contrast, resulted in a substantially smaller probability of a level of performance of over 90% (although the percentage of almost perfect solution was non-negligible here, it was 20–40%).

(6) Among students arriving at perfect solutions, the most frequent strategy was the most efficient, completely odometrical strategy for all three test items in both age groups (with a

frequency of over 50%). This was followed in frequency by the nearly odometrical strategy (20–25%) and the least frequent strategy among perfect solutions was random enumeration (10–15%). As was predicted, perfect solutions were therefore more likely to be generated by participants using the more efficient strategies but the most rudimentary, random strategy could also lead to a perfect answer (H_5).

(7) The strategies used for the solution of the three problems correlates with each other in both age groups ($p<0.01$). Adjacent pairs of items (Items 1 and 2 and Items 2 and 3) were typically characterised by identical strategies (40–50%), which was followed in frequency by solutions using progressively less complex strategies (around 30%) and finally by solutions using progressively more complex strategies (around 25%). That is, as problem solvers moved forward in the test, all three patterns of change in strategy use occured (H_6).

(8) Looking at the three items in combination, we found that about 60% of students followed the same strategy in solving at least two of the problems and about a quarter of participants used the same strategy in solving all three of the problems. In the latter group, the highest percentage of problem solvers (about 60%) used the completely odometrical strategy. They were followed in frequency by those using the random strategy to enumerate their solutions (35 and 25%).

RELEVANCE OF THE RESEARCH PROGRAMME

The research programme of the study of enumerative combinatorial problems described in the dissertation defined three overall aims – the development of a measurement instrument, the analysis of the comprehension of combinatorial problems and the study of strategy use, – which were achieved by conducting the four studies presented above. The following paragraphs will outline in what ways the results of the research programme enrich our knowledge of combinatorial reasoning from the perspective of pedagogical and education science.

The measurement instrument we developed is a valuable addition to the body of online measurement instruments suitable for the assessment of combinatorial reasoning. In our view, an advantage of the measurement instrument is that the items included adhere to a completely uniform structure while the stories in which the problems are embedded match the criteria of the given operation and vary across items. The test section of the measurement instrument is suitable for the educational measurement of performance but in its current form, the assistance of a researcher is required in order to analyse the details of solving the problems. If the automatically computed values of the variables defined in the research programme were built into the eDia platform, they would be accessible to teachers on completion of the test and they could provide useful information for the diagnostic assessment of the solutions of combinatorial problems in a classroom environment.

We have developed a further method of analysing solutions to enumerative combinatorial problems that provides a detailed picture of errors occurring during the completion of the task. The variables we have defined can be used to evaluate students' comprehension of the combinatorial problem with respect to the three criteria of combinatorial operations: set size, replacement and orderedness. The development of the evaluation procedure resulted in a set of codes that allow the automatic computation of the values of the variables for the items included in the test (and for other items of the same structure). This means that provided that technology-based testing is employed, the values of the variables will immediately be at our disposal. The use of the method proposed here provides – in addition to performance – valuable information about the possible dimensions of error for the evaluation of students' answers when a non-perfect answer is given. The information thus gained can even be used at a later point to foster reasoning skills since it indicates what difficulties each student encountered with each criterion of each operation. Our analyses of the functioning of the variables enriches our knowledge about combinatorial reasoning among 4th and 6th graders.

In order to be able to analyse strategy use, new event query functions have been added to the eDia online measurement system with the assistance of software developers. These functions allow us to reconstruct the process of solving the problem and may also provide valuable information for researchers wishing to study areas other than the solution of problems assessing combinatorial reasoning.

To our knowledge, ours was the first study to investigate strategies used during the solving of enumerative combinatorial problems. We are also unaware of research findings obtained abroad from a large-scale study with computerised data collection and using log files to analyse combinatorial strategies. Our results on strategy use further our understanding of the nature of the process of solving problems involving the Cartesian product of sets among 4th and 6th graders.

REFERENCES USED IN THE THESIS BOOKLET (WITHOUT OWN PUBLICATIONS)

Adey, P., & Csapó, B. (2012). A természettudományos gondolkodásfejlesztése és értékelése. In B. Csapó & G. Szabó (Eds.), *Tartalmi keretek a természettudomány diagnosztikus értékeléséhez* (pp. 17–58). Budapest: Nemzeti Tankönyvkiadó.

Batanero, C., Godino J. D., & Navarro-Pelayo, V. (1997). Combinatorial reasoning and its assessment. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 239–252). Amsterdam: IOS Press.

Bräuning, K. (2019). Development of strategies for a combinatorial task by a 5 year old child. Eleventh Congress of the European Society for Research in Mathematics Education (CERME11), Utrecht University, Feb 2019, Utrecht, Netherlands. hal-02414857

Csapó, B. (1988). *A kombinatív képesség struktúrája és fejlődése.* Budapest: Akadémiai Kiadó.

Csapó, B. (2001). A kombinatív képesség fejlődésének elemzése országos reprezentatív felmérés alapján. *Magyar Pedagógia, 101*(4), 511–530.

Csapó, B. (2003). *A képességek fejlődése és iskolai fejlesztése.* Budapest: Akadémiai Kiadó.

Csapó, B., Csíkos, Cs., & Molnár, Gy. (2015, Eds.). *A matematikai tudás online diagnosztikus értékelésének tartalmi keretei*. Budapest: Oktatáskutató és Fejlesztő Intézet.

Csapó, B., & Pásztor, A. (2015). A kombinatív képesség fejlődésének mérése online tesztekkel. In B. Csapó & A. Zsolnai (Eds.), *Online diagnosztikus mérések az iskola kezdő szakaszában* (pp. 367–386). Budapest: Oktatáskutató és Fejlesztő Intézet.

DeTemple, D., & Webb, W. (2014). *Combinatorial reasoning. An introduction to the art of counting*. Hoboken, New Jersey: John Wiley & Sons, Inc.

English, L. D. (1991). Young children's combinatoric strategies. *Educational Studies in Mathematics, 22*(5), 451–474.

English, L. D. (1993). Children's strategies for solving two- and three-dimensional combinatorial problems. *Journal for Research in Mathematics Education, 24*(3), 255–273.

English, L. D. (2005). Combinatorics and the Development of Children's Combinatorial Reasoning. In G. A. Jones (Ed.), *Exploring probability in schools: Challenges for teaching and learning* (pp. 121–141). Dordrecht: Kluwer.

English, L. D. (2016). Revealing and capitalising on young children's mathematical potential. *ZDM Mathematics Education, 48*(7), 1079–1087.

Hadar, N., & Hadass, R. (1981). The road to solving a combinatorial problem is strewn with pitfalls. *Educational Studies in Mathematics, 12*(4), 435–443.

Hajduné Holló, K. (2004). Az elemi kombinatív képesség fejlődésének kritériumorientált diagnosztikus feltárása 4–8 évesek körében. *Magyar Pedagógiai, 104*(3), 263–292.

Halani, A. (2012). Students' ways of thinking about enumerative combinatorics solution sets: the odometer category. In S. Brown, S. Larsen, K. Marrongelle, & M. Oehrtman (Eds.), *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education* (pp. 59–68). Portland, Oregon: The Special Interest Group of the Mathematics Association of America (SIGMAA) for Research in Undergraduate Mathematics Education.

Höveler, K. (2018). Children's Combinatorial Counting Strategies and their Relationship to Conventional Mathematical Counting Principles. In E. W. Hart & J. Sandefur (Eds.), *Teaching and Learning Discrete Mathematics Worldwide: Curriculum and Research. ICME-13 Monographs* (pp. 81–92). Springer International Publishing.

Inhelder, B., & Piaget, J. (1967). *A gyermek logikájától az ifjú logikájáig.* Budapest: Akadémiai Kiadó.

Lockwood, E. (2013). A model of students' combinatorial thinking. *The Journal of Mathematical Behavior, 32*(2), 251–265.

Lockwood, E. (2015). The Strategy of Solving Smaller, Similar Problems in the Context of Combinatorial Enumeration. *International Journal of Research in Undergraduate Mathematics Education, 1*(1), 339–362.

Mashiach-Eizenberg M., & Zaslavsky, O. (2004). Students' verification strategies for combinatorial problems. *Mathematical Thinking and Learning, 6*(1), 15–36.

Melusova, J., & Vidermanova, K. (2015). Upper-secondary Students' Strategies for Solving Combinatorial Problems. *Procedia – Social and Behavioral Sciences, 197*, 1703–1709.

Molnár, Gy. (2017). Az interaktív problémamegoldó képesség fejlettségi szintjét befolyásoló képességek. In J. Kerülő, T. Jenei, & I. Gyarmati (Eds.), *Program és absztrakt kötet. XVII. Országos Neveléstudományi Konferencia* (pp. 464). Nyíregyháza.

Molnár, Gy., & Csapó, B. (2019). A diagnosztikus mérési rendszer technológiai keretei: Az eDia online platform. *Iskolakultúra, 29*(4–5). 16–32.

Mwamwenda, T. S. (1999). Undergraduate and graduate students' combinatorial reasoning and formal operations. *Journal of Genetic Psychology, 160*(4), 503–505.

Nagy, J. (2004). Az elemi kombinatív képesség kialakulásának kritériumorientált diagnosztikus feltárása. *Iskolakultúra, 14*(8), 3–20.

Palmér, H., & van Bommel, J. (2018). The role of and connection between systematization and representation when young children work on a combinatorial task. *European Early Childhood Education Research Journal, 26*(4), 562–573.

Pap-Szigeti, R. (2009). *Kritériumorientált képességfejlesztés tantárgyi tartalmakkal az 5. évfolyamon*. PhD értekezés, Szeged.

Piaget, J. (1970). *Válogatott tanulmányok*. Budapest: Gondolat Kiadó.

Poddiakov, A. (2011). Didactic objects for development of young children's combinatorial experimentation and causal-experimental thought. *International Journal of Early Years Education, 19*(1), 65–78.

Scardamalia, M. (1977). Information Processing Capacity and the Problem of Horizontal "Décalage": A Demonstration Using Combinatorial Reasoning Tasks. *Child Development, 48*(1), 28–37.

Shin, J., & Steffe, L. P. (2009). *Seventh-graders' use of additive and multiplicative reasoning for enumerative combinatorics problems*. In S. L. Swars, D. W. Stinson, & S. Lemons-Smith (Eds.), *Proceedings of the 31st Annual meeting of the North American Chapter of the International Group* *for the Psychology of Mathematics Education* (pp. 170–177). Atlanta, GA: Georgia State University.

Szitányi, J., & Csíkos, Cs. (2015). Performance and strategy use in combinatorial reasoning among pre-service elementary teachers. In K. Beswick, T. Muir, & J. Wells (Eds.), *Proceedings of the 39th Conference of the International Group for the Psychology of Mathematics Education* (pp. 4225–4232). Hobart: International Group for the Psychology of Mathematics Education.

Tillema, E. S. (2013). A power meaning of multiplication: Three eighth graders' solutions of Cartesian product problems. *The Journal of Mathematical Behavior, 32*(3), 331–352.

Wu, H., & Molnár, G. (2018). Interactive problem solving: Assessment and relations to combinatirial and inductive reasoning. *Journal of Psychological and Educational Research, 26*(1), 90–105.

Zentai, G., Hajduné Holló, K., & Józsa, K. (2018). Új mérőeszközök a gondolkodás vizsgálatára 4–8 éves korban. In O. Endrődy-Nagy & A. Fehérvári (Eds.), *HERA Évkönyv 2017: Innováció, kutatás, pedagógusok* (pp. 175–189). Budapest: Magyar Nevelés- és Oktatáskutatók Egyesülete.

OWN PUBLICATIONS CITED IN THE DISSERTATION

Gál-Szabó, Zs. (2019). Felsoroló kombinatív problémák megoldása során használt stratégiák mérésének előkészítése. *Neveléstudomány: Oktatás – Kutatás – Innováció, 7*(1) 31–46.

Gál-Szabó, Zs., & Bede-Fazekas, Á. (2020). Formalization of Odometer Thinking and Indices for the Classification of Combinatorial Strategies. *International Electronic Journal of Mathematics Education, 15*(1), em0546.

Gál-Szabó, Zs., & Korom, E. (2018a). A teljesítmények és a feladatmegoldásról alkotott tanulói vélemények összefüggése a kombinatív gondolkodás mérése kapcsán 4. és 6. évfolyamon. In T. Vidákovich & N. Fűz (Eds.), *PÉK 2018. XVI. Pedagógiai Értékelési Konferencia. Program és összefoglalók* (p. 36). Szeged: Szegedi Tudományegyetem.

Gál-Szabó, Zs., & Korom, E. (2018b). Combinatorial reasoning among 4th and 6th grade pupils: preliminary results of a large-scale study. In T. Vidákovich & N. Fűz (Eds.), *PÉK 2018. XVI. Pedagógiai Értékelési Konferencia. Program és összefoglalók* (p. 37). Szeged: Szegedi Tudományegyetem.

Gál-Szabó, Zs., & Korom, E. (2018c). Felsoroló kombinatív feladatok megértésének vizsgálata az elemszám, az ismétlődés és a felcserélhetőség kritériumok alapján. *Magyar Pedagógia, 118*(4), 385–413.

Gál-Szabó, Zs., & Korom, E. (2019). A kombinatív gondolkodás longitudinális vizsgálata: a teszten nyújtott teljesítmény és a feladatok megértésének változása 4–5. és 6–7. évfolyamok között. *Magyar Pedagógia, 119*(1), 3−18.

Gál-Szabó, Zs., Korom, E., & Steklács, J. (2019). Feladatmegoldó viselkedés és kombinatív stratégiák felsoroló kombinatív probléma megoldása során. In J. Steklács (Ed.), *Szemkamerás vizsgálatok a pedagógiai kutatásban* (pp. 25–50). Kaposvár: Kaposvári Egyetem Pedagógiai Kar.

Szabó, Zs. G. (2015). *A kombinatív képesség rövid távú fejleszthetőségének vizsgálata 3. évfolyamon*. Szakdolgozat. Szeged.

Szabó, Zs. G., & Korom, E. (2016). A feladatmegoldási idő vizsgálata kombinatív képességet mérő teszten 3. évfolyamosok körében. In Gy. Molnár & E. Bús (Eds.), *PÉK 2016. XIV. Pedagógiai Értékelési Konferencia. Program – Előadás-összefoglalók* (pp. 149). Szeged: Szegedi Tudományegyetem.

Szabó, Zs. G., & Korom, E. (2017). Kombinatív stratégiák feltárására fejlesztett mérőeszköz és kipróbálásának előzetes eredményei. In J. Kerülő, T. Jenei, & I. Gyarmati (Eds.), *XVII. Országos Neveléstudományi Konferencia. Program és absztrakt kötet* (p. 510). Nyíregyháza: MTA Pedagógiai Tudományos Bizottság, Nyíregyházi Egyetem.

Szabó, Zs. G., Korom, E., & Pásztor, A. (2015). A kombinatív képesség rövid távú fejleszthetősége 3. évfolyamon természettudományos kontextusban. *Magyar Pedagógia, 115*(4), 383–401.