Thinking Skills and Creativity 50 (2023) 101408

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc

Understanding transitions in complex problem-solving: Why we succeed and where we fail

Gyöngyvér Molnár [a,*], Samuel Greiff [b]

[a] Institute of Education, University of Szeged, MTA–SZTE Digital Learning Technologies Research Group, Hungary
[b] Computer-Based Assessment Research Group, Department of Behavioural and Cognitive Sciences, University of Luxembourg, Luxembourg

ARTICLE INFO

Keywords:
Complex problem-solving
Knowledge transition
Latent transition profiles
Time-on-task
Number of clicks

ABSTRACT

Problem-solving is an essential transversal and learning skill in the 21st century. To better support it, we need a deeper understanding of its processes, i.e., problem exploration, problem representation and knowledge use. Understanding the mechanism behind the transitions between these processes plays an important role in supporting its development. With a sample of $N = 1828$ first-year university students, this study is the first to measure how students transition between the different problem-solving processes and how this influences their overall problem-solving performance as well as how this is related to their other test-taking behaviours. Results indicate that mastering the first transition – correctly understanding and depicting the problem structure – plays a crucial role in problem-solving. Students who failed to master it regularly failed in the next transition. Problem complexity strongly influenced transitions during the entire problem-solving process. Based on students' behavioural patterns, we distinguished four qualitatively different latent transition classes: expert transitioners, advanced transitioners, beginner transitioners and non-transitioners. The number of interactions proved to be a more effective profile characteristic, especially on high-complexity problems, than time spent on the problem-solving process. The results of the current study provide important insights into how students' transition between the different problem-solving processes and how their test-taking behaviour indicates their transition skills.

1. Introduction

Education today does not focus solely on teaching specific academic disciplines. It also aims to foster essential transversal skills via the incorporation of educational technology and to prepare pupils for the future, for the unknown and for work as well as to solve problems in uncertain situations. In problem-solving, students have to formulate and analyse the problem by identifying relevant elements and facts in a given scenario (Hmelo-Silver, 2004). This aids in establishing a representation of the problem. As students' understanding of the problem scenario grows, they start generating hypotheses about possible solutions and the specific sequence of actions needed to reach this solution. During this process, they identify knowledge deficiencies tied to the problem. These knowledge deficiencies form the basis of the learning process that students should research and acquire during their self-directed, enquiry-based learning.

To better support students' learning processes and learning outcomes, we need a deeper understanding of their thinking processes,

* Corresponding author.
E-mail address: gymolnar@edpsy.u-szeged.hu (G. Molnár).

https://doi.org/10.1016/j.tsc.2023.101408
Received 6 January 2023; Received in revised form 24 September 2023; Accepted 1 October 2023
Available online 2 October 2023
1871-1871/

particularly when it comes to problem-solving. To this end, researchers have started applying the tools of learning analytics to information gathered and logged (Csapó et al., 2012) during the learning and assessment process (Zhang et al., 2019). Analysis of this information can contribute to a better understanding of the phenomenon under examination and help to explain the mechanism behind the score values of the variables observed, especially students' thinking processes while learning and solving problems. For example, analysing students' task completion and test-taking behaviour may inform us about where and when they have made mistakes and what type of mistake they have made (Nicolay et al., 2021; Stadler et al., 2020), which behaviour pattern results in success or failure (Eichmann et al., 2020a, b; Molnár & Csapó, 2018), how much learners engage in collaborative learning (Nasir et al., 2021), what kind of exploration strategy they use during the problem-solving process (Molnár et al., 2022) and, based on this, how many qualitatively different exploration profiles can be distiguished (Eichmann et al., 2020a; Greiff et al., 2016, 2018; Molnár et al., 2022; Mustafić et al., 2019; Wu & Molnár, 2021), how many interactions students engage in (Greiff et al., 2016; Molnár et al., 2022) and how much time they spend in the problem-solving process (Goldhammer et al., 2014). Researchers have also examined how cognitive factors, such as subject matter knowledge, and affective factors, such as engagement, influence students' strategy efficacy and strategy selection (e.g., Newton et al., 2020) or pause time before problem-solving predicts effective strategy (Chan et al., 2022).

Problem-solving and complex problem-solving represent a step-by-step process. Three steps, or processes, have been identified in the process of complex problem-solving: (1) problem exploration, (2) problem representation and (3) knowledge use (Fischer et al., 2012; Molnár & Csapó, 2018). In the present study, problem exploration is operationalized as exploring a system to generate knowledge about a problem space by interacting with a problem scenario (i.e., engaging in trial and error). Problem representation is operationalized as building mental maps, and knowledge use is operationalized as using and putting the newly acquired knowledge into practice by managing learning goals.

This study assumes that the transitions between these three problem-solving processes play an important role in understanding the overarching process of problem-solving. Beyond Nicolay et al. (2021), who focused on the transition from knowledge acquisition to knowledge application, no research has thus far examined the transitions that students encounter when working on complex problems, largely since these transitions are not directly visible in behaviour. With this gap in mind, the present study investigates two transitions and how success or failure with these transitions influences problem-solving and builds empirically distinguishable, qualitatively different latent profiles of students as they work on ten fictitious complex problems of increasing difficulty built within the MicroDYN approach. The second part of the paper focuses on the validation of the results in understanding the mechanism behind the two transitions and involves other process indicators of test-taking behaviour in the analysis, such as time-on-task and number of clicks.

1.1. Understanding transitions in complex problem-solving

As noted above, three processes are distinguished in solving complex problems: (1) problem exploration, (2) problem representation and (3) knowledge use (Fischer et al., 2012; Greiff et al., 2012; Molnár & Csapó, 2018). Problem exploration and problem representation are involved in the knowledge acquisition phase of the problem-solving process, while knowledge use, that is, finding a solution to a specific problem, defines the knowledge application phase as the second key problem-solving phase beyond knowledge acquisition (Greiff et al., 2013). We assume that by understanding the mechanism behind transitions between these three, empirically distinguished processes (i.e., the transition between (1) problem exploration and problem representation and that between (2) problem representation and knowledge use), we can advance our knowledge of students' thinking processes in solving complex problems. With this gap in mind, the present study examines two transitions and how success or failure with them influences problem-solving.

We examine the first transition students encounter between problem exploration and problem representation and the second one they meet between correct problem representation and knowledge use, independently of the first transition. Specifically, we go beyond monitoring problem-solving achievement as a single indicator and address the mechanism behind successful problem-solving and learning in uncertain situations by focusing on knowledge transition – from collecting relevant information via hypothesis formation to problem representation (first transition) and from problem representation to knowledge use (second transition) – in problem-based environments. Participants proved to be successful in the first transition if they were able to apply a theoretically effective exploration strategy to generate knowledge about the problem space and if they were able to interpret and draw a cognitive representation of this mental map in the form of a concept map. Subsequently, participants proved to be successful in the second transition if they were able to put previously acquired knowledge from their mental model into practice (and/or interpret the right mental model presented on screen) by executing effective goal-directed manipulations to reach certain predefined goals.

Nicolay et al. (2021) addressed the question of transitions using traditional achievement data between the two major phases of complex problem-solving: knowledge acquisition and knowledge application. They argued that a great proportion of students fails to engage in successful knowledge transition, whose rates vary according to item complexity. We go deeper and examine both of the transitions between the three complex problem-solving processes (problem exploration, problem representation and knowledge use) by analysing behavioural logfile data, and we investigate how success or failure with these transitions influences problem-solving.

1.2. Test-taking behaviour might be a valid indicator of the construct under investigation

Cognitive assessment provides information not only about a student's problem-solving skills and functional capacities (Heinonen et al., 2011), but also about their test-taking behaviour (Meyer et al., 2001), which affects their cognitive test performance (Heinonen et al., 2011) as well as having potential implications for further academic outcomes (Blair & Razza, 2007). "If test-taking behaviour actually represents specific cognitive and metacognitive processes, individual differences in ability to solve a task should be reflected in task behaviour" (Stadler et al., 2020. p. 1). That is, measuring test performance with behavioural observations provides more valuable information in planning interventions, while behavioural data may provide information about future cognitive achievement.

Time-on-task, that is, time spent providing an answer, is amongst the most investigated behavioural data; it is a major characteristic of the task completion process which underlies different behaviours (Goldhammer et al., 2014). According to early studies, longer time relates to better performance (Schiffring & Schneider, 1977), and simpler tasks which allow the application of routine cognitive processing are related to shorter time; at the same time, more complex tasks allow for multiple ways of arriving at the same solutions (Stadler et al., 2020). This has also been confirmed by Alzoubi et al., (2013) and Eichmann et al. (2020a), who stated that more time allows participants longer planning and better planned solutions, which resulted in significantly higher achievement. In contrast, Greiff et al. (2016) argued that spending too much time on a task is associated with poor performance and suggested a U-shaped relation. Based on recent research results, Chan et al. (2022) interpreted extra time as a proxy indicator of thinking before problem-solving. They observed that pause time – including longer time-on-task – "may be an indicator of student thinking before problem-solving, and provide insights into using data from online learning platforms to examine students' problem-solving processes" (Chan et al., 2022, p. 1). As a result of contradictory results obtained in the different cognitive domains, Naumann & Goldhammer (2017) suggested that time-on-task has no uniform interpretation but is a function of task difficulty and individual skill.

Beyond time-on-task, number of clicks or number of interactions on an assessment task also supplies important information about students' engagement, involvement and thinking processes, although less attention has been paid to this issue. Goldhammer et al. (2014) distinguished between interactions that aim to access information and those focused on the use of the information accessed. As regards problems requiring controlled processing, he and his colleagues concluded that low-achieving students typically engage in fewer interactions than high achievers. Eichmann et al. (2020a) confirmed the positive correlation between problem-solving achievement and number of interactions, while Lotz et al. (2017) and Stadler et al. (2020) reported on negative correlations between number of interactions and GPA. They argued that students with greater ability appear to spend more time while needing to engage in fewer interactions for successful problem-solving (Stadler et al., 2020). Based on all these research results that highlight the predictive power of process data on students' achievement, we applied these variables as validation variables for results obtained on students' transition skills.

1.3. The present study

The present study builds on the assumption that problem-solving is a step-by-step process that consists of (1) problem exploration, (2) problem representation and (3) knowledge use. The study also assumes that the transitions between these three processes (i.e., the transition between (1) problem exploration and problem representation and that between (2) problem representation and knowledge use) play an important role in understanding the overarching process of problem-solving. We also build on the research results, which suggest that behavioural indicators like time-on-task and number of clicks may also be good indicators of students' success, which is strongly influenced by their transition skills.

To our knowledge, this study is the first to measure how students' transition between the different problem-solving processes and how this influences their overall problem-solving performance as well as how this is related to their other test-taking behaviours. Thus, we also used learning analytics to examine questions about the relation of test-taking behaviour to students' latent class transition profiles investigated by sequential transition behaviour patterns.

The goal of our study is to answer the following research questions (RQs):

RQ1: Describing the transitions
RQ1a: What is the proportion of first-year university students who master both, only the first, only the second or neither of the transitions in the problem-solving process?
RQ1b: How does the number of successful transitions vary with the complexity of the problem?
RQ1c: How many and what kind of latent class profiles can be distinguished based on students' transition behaviour pattern analyses on the test level?
RQ2: The relation of the transitions to test-taking behaviour
*RQ2a: How does students' test-taking behaviour (*i.e.*, time-on-task and number of interactions) differ depending on the number of successful transitions (theory-driven, problem-level transition-focused approach)?*
*RQ2b: How do students' latent transition profiles differ with regard to their real problem-solving behaviour characteristics (*i.e.*, time-on-task and number of interactions; latent transition profile-based approach)?*

2. Research methods

2.1. Participants

Participants were students starting their studies at one of the largest and highest-ranked universities in Hungary. The university has twelve faculties (e.g., humanities and social sciences, science, medicine, law and economics), and students from all the faculties were involved in the assessment. A total of 1844 students, that is, 44.8 % of the target population, participated in the study (age mean = 19.8; SD = 1.74), 59.8 % of them being female. After data cleaning, that is, deleting all the students who did not reach the end of the test (less than 1 % of the sample), 1828 students remained in the sample. Students' participation was voluntary, and they received one

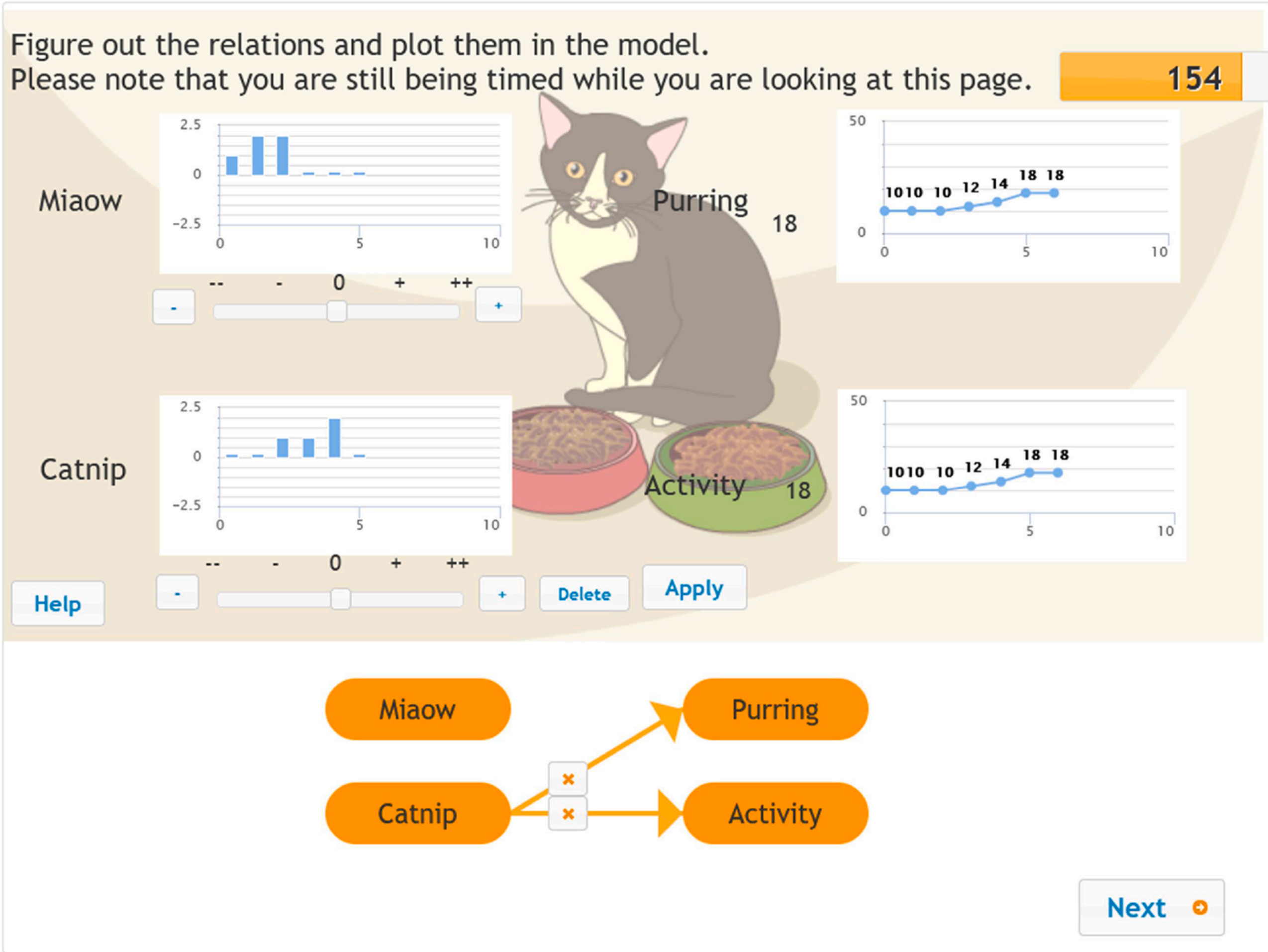

Fig. 1. Screenshot of the MicroDYN task “Cat” Phase 1 – knowledge acquisition (processes: problem exploration and model building). The controllers of the input variables range from “- -” (value = -2) to “++” (value = +2). The model representation is shown at the bottom of the figure.

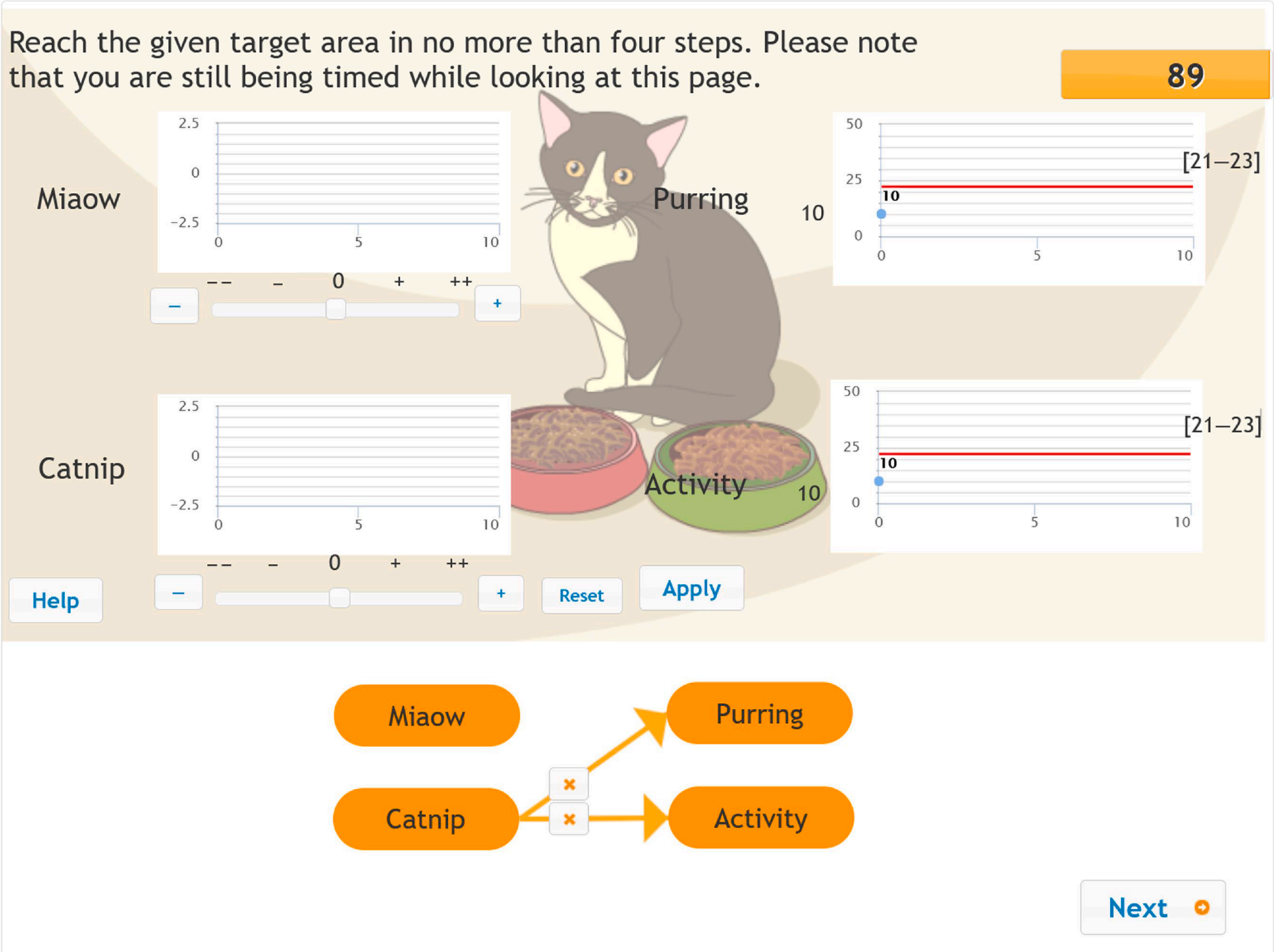

Fig. 2. Screenshot of the MicroDYN task "Cat" Phase 2 – knowledge application (process: knowledge use). The controllers of the input variables range from "- -" (value = -2) to "++" (value = +2). The correct representation of the mental model is shown at the bottom of the figure.

course credit for completion of the tests as an incentive.

2.2. Instrument

To test students' problem-solving skills, specifically, their transition skills and the mechanism behind them, including the facets of knowledge acquisition and knowledge application, we searched for problems where the complexity of the problem can be defined, scaled and distinguished and where students' domain-specific knowledge plays as little of a role as possible. Based on existing research results on problem-solving, we decided on a widely used and validated computer-based instrument for complex problem-solving (CPS), MicroDYN (Funke, 2001; Greiff et al., 2012). The computer-simulated problems contain up to six interrelated variables in uncertain problem situations. They capture three overarching processes of problem-solving: problem exploration and problem representation (in the first phase, labelled knowledge acquisition) and use of knowledge (in the second phase, labelled knowledge application).

In the process of exploring the CPS problems (in phase 1), students were expected to generate knowledge about the elements and their relationships in the given problem space by interacting with the problem scenario. Effective interactions with the problem space formed the foundation of successful problem presentation. To do this, students first had to identify the type of variable (input or output). Second, they had to manipulate the values for the input variables – as many times as they liked within 180 s – and then test, identify and interpret the changes in the output variables to increase their understanding of the problem space (see Fig. 1). For example, in the first phase of a problem called "Cat" (see Fig. 1), students were expected to adjust the levels of two input variables, which represented fictitious cat foods ("Miaow" and "Catnip") and observe changes in the output variables, which represented the cats' level of movement and purring. Students were asked to adjust the level of the cat foods and their combinations in any way to discover their impact on the cat's movement and purring. To test the different settings for the amount and combinations of cat food, they were expected to click the Apply button. As students' knowledge of the problem scenario grows, they can start building a mental model of the structure of the problem and identify knowledge deficiencies tied to the problem, which requires additional interaction with and learning about the problem space. After exploring, they were asked to present the cognitive representation of the newly built mental map in a model presented on screen. To do this, they were expected to draw the detected relations in the form of arrows between the variables presented on screen (Fig. 1), that is, in the given example, between the different cat foods (Miaow and Catnip) and the cat's different behaviours (purring and other activity). Thus, in the first phase, two problem-solving processes, problem exploration and problem representation, are measured.

In the second phase (see Fig. 2) of the problem-solving process, students build the correct representation of the mental model (the knowledge they had to explore and acquire in the first phase of the problem-solving process), and they apply it by managing the learning goals and reach the given target values of the output variables within a given time frame (90 s) in no more than four combinations of input variable settings. In the case of the Cat problem, students received the correct relation between cat food and movement/purring in the form of the correct concept map, and they are expected to change the levels of two different cat foods to reach the target values given for movement and purring in no more than four steps (clicking four times on the Apply button). Thus, in the second phase, one additional problem-solving process, knowledge use, is measured. Based on empirical evidence (see, e.g., Greiff et al., 2013; Molnár et al., 2021; Molnár & Csapó, 2017; Schweitzer et al., 2013), these scenarios are good measures of different phases of the whole problem-solving process.

In total, the CPS test consisted of ten problems with fictitious cover stories and increasing item complexity. CPS items can be described with six item characteristics (Stadler et al., 2016): (1) the number of input and (2) output variables and (3) the number and (4) type of relations (i.e., direct, autonomous) and the number of (5) irrelevant input (i.e., manipulating these variables has no impact on the system) and (6) irrelevant output variables (not related to any input variables). In the case of the Cat problem, both the number of input and output variables and the number of relations are two, the type of relation is direct, and the problem scenario contains one irrelevant input variable. With regard to item complexity and based on earlier studies (Molnár et al., 2022; Nicolay et al., 2021) on the number and type of variables and relations, we distinguished between low-, medium- and high-complexity problems. That is, problem-solving tasks became more complex with the growing number of variables and relations and with the use of autonomous changes beyond direct relations. MicroDYN problems of low complexity (two problems on the test) involved two input and two output variables with two relations (direct effect with no autonomous changes in variables – such as the "Cat" problem in Figs. 1 and 2). Medium-complexity items (four problems on the test) consisted of five or six variables altogether (three input and two or three output variables) with three or four direct effects. The difference between medium- and high-complexity items (four problems on the test) was that the latter also contained indirect effects (Greiff et al., 2013), that is, autonomous changes in output variables without changing the values of the input variables beyond the direct relations. Irrelevant variables were also typical of high-complexity items. Appendix 1 provides examples of structural relations and equations for low-, medium- and high-complexity problems.

The reliability of the test was good. Based on dichotomous scoring, that is, the traditional CPS indicators (Greiff et al., 2013) for phases 1 and 2 (phase 1 for knowledge acquisition and phase 2 for knowledge application), reliability proved to be $\alpha = 0.882$; using the process-oriented approach and splitting phase 1 into two distinguishable processes, Cronbach's alpha proved to be 0.944 for problem exploration, 0.858 for problem representation and 0.758 for knowledge use using binary scores.

2.3. Procedures

2.3.1. Data collection

The data collection was carried out in a large computer room at the university learning and information centre during the first four weeks of the semester. The test was administered using the eDia online platform (Csapó & Molnár, 2019). Students had 60 min to finish it and the related background questionnaire. The whole research project involved two testing sessions of two hours each consisting of measures of problem-solving, mathematical reasoning, reading comprehension, working memory and an additional domain, which was chosen by the dean of each of the faculties (e.g., ICT literacy, learning strategies, inductive reasoning, financial literacy, social sciences and science). At the beginning of the test, participants were provided instructions on the user interface, including a warm-up task. Students received immediate feedback on their average achievement after completing the test and additional, detailed feedback with normative comparative data on their performance a week after the testing was closed.

According to the national and institutional guidelines, ethical approval was not required for this study. The assessments which provided data for this study were integrated parts of the educational processes of the participating university. Participation was voluntary, and participants received a course credit for active participation. All of the students in the assessment had turned 18 by the time of the assessment. That is, it was not required or possible to request and obtain written informed parental consent from the participants, but all of them confirmed with their signature that their data could be used for educational and research purposes at both the faculty and university levels.

2.3.2. Dataset

We built a dataset consisting of the scored answer data for the three processes, problem exploration, problem representation and knowledge use, and the log data (interaction data, which describe the manipulation behaviour of the students, such as number of interactions and time-on-task). All these actions are nested on the individual level, as each student was expected to solve ten CPS problems as well as provide data for the three processes and log data for each problem scenario. That is, each student has multiple observations that are not independent but rather correlated because they are associated with the same students, thus offering the possibility to go beyond a single observation and learn about student's transition behaviour in multiple cases and in a learning environment.

2.3.3. Scoring and labelling the log data

In the first problem-solving phase during the first CPS process, we scored the effectiveness of problem exploration, that is, the interaction behaviour of the students, based on the collected logfiles. In order to map and describe the exploration strategy of the students, we used the labelling procedure developed by Molnár and Csapó (2018), which has also been used in international assessments (e.g., Molnár, 2021, 2022; Wu & Molnár, 2021). If the interaction behaviour provided all the information about the detected relations, it was considered as a theoretically effective strategy and assigned a score of 1; otherwise, the manipulation behaviour was considered ineffective, and students earned a score of 0. For example, on the "Cat" problem, if the participant tested the effect of the cat foods on the cat's behaviour by manipulating both cat foods (both input variables) at the same time (that is, feeding the cat a mixture of the two foods), thus finishing the first part of the problem, they received a score of 0. This type of manipulation cannot provide sufficient information on the separate effect of each of the cat foods on the cat's behaviour (described by the output variables). However, if participants first tested the effect of Miaow on the cat's behaviour and then separately tested the effect of Catnip, they received a score of 1. This is because this type of manipulation proved all the information about the effects of the cat foods on the cat's behaviour. Across all ten problems, each student received ten binary (correct/incorrect) scores on their problem exploration based on the strategies they applied (for a more detailed description, see Molnár & Csapó, 2018).

In addition to this and also taking place during the first phase but related to the second process, we scored the visualized cognitive representation of their mental map of the problem space, which indicated the detected relations in the form of arrows between the variables presented on screen (see Fig. 1). A completely matching problem structure was assigned a score of 1 (e.g., on the "Cat" problem, if students draw one arrow from the Catnip cat food to purring and another arrow from Catnip to movement but no other arrows); otherwise, the response was considered incorrect and earned a score of 0. Thus, across all ten problems, each student received ten binary (correct/incorrect) scores on their visualized cognitive mental map representation of the problem space.

Finally, in the second phase, after receiving the correct representation of the mental model, we scored the third CPS process, students' knowledge use. The answers were marked correct ("1") if students managed to reach the given target values of the output variables within the given constraints (e.g., within a pre-specified number of steps; in the example, if students managed to increase the level of the cat's movement and purring from 10 to 21–23); otherwise, the solution was considered incorrect and assigned a score of 0. Thus, across all ten problems, each student received ten binary (correct/incorrect) scores on their knowledge use.

Beyond the three processes outlined above, log data indicating time-on-task and number of interactions were marked separately for each problem and for each student. Thus, in total, five measures were available for each problem targeted by each student: problem exploration (binary), problem representation (binary), knowledge use (binary), time-on-task (continuous) and number of interactions (continuous). The latter two were used to investigate differences in other variables related to the different combinations of successful and unsuccessful transitions.

2.3.4. Analysis

The main goal of this paper is to examine the transitions between the three processes, that is, how the students transition (1) from problem exploration to problem representation and (2) from problem representation to knowledge use. To test the transition of the amount of extracted information (use of a theoretically right or wrong strategy) to problem representation and from the right problem representation or from the right exploration to its direct or indirect use, we used a decision tree-based approach (see Fig. 2). In this, we can form four artificial groups of students on the problem level: (1) students who succeeded in both transitions (i.e., they were successful in the transitions from problem exploration to problem representation and from problem representation to knowledge use) (Group A; score: 1 – 1 – 1); (2) students with partly successful transitions who succeeded in the first but failed in the second (i.e., they were only successful in the transition from problem exploration to problem representation) (Group B; score: 1– 1 – 0); (3) students with partly successful transitions who failed in the first but succeeded in the second (i.e., they were only successful in the transition from problem representation to knowledge use) (Group C; score: 0 – 1 – 1); and, finally, (4) students with no successful transitions (i.e., they were successful in neither of the transitions) (Group D; score: different patterns: 1 – 0 –1, 1 – 0 – 0, 0 – 1 – 0, 0 – 0 – 1, 0 – 0 – 0). However, Group D proved to be a heterogenous group and was split up further into five different sub-groups using the decision tree approach (see Fig. 2):

- Students in Sub-group D1 applied the right exploration strategy but failed to build and visualize their mental map on the concept map presented on screen. After building the proper concept map, which they needed to do to solve the problem, they were able to

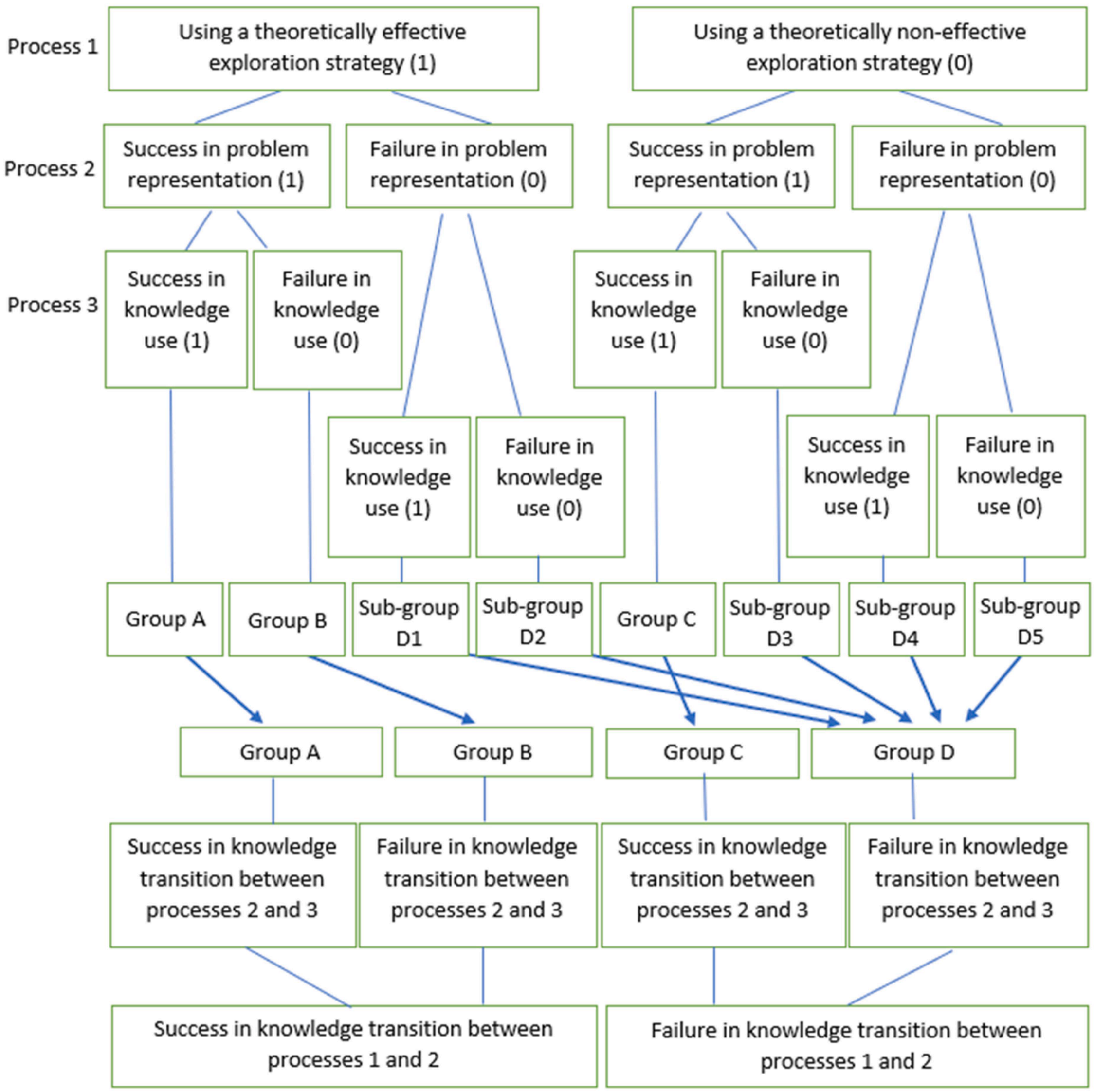

Fig. 3. Integrated model: decision tree with the different categories of students.

transfer the knowledge they had retrieved from the first to the third phase of the problem-solving process (score: 1 – 0 – 1). They were partly successful in the transition but in a different way than students in Group B or C. They managed an indirect transition between the first (problem exploration) and third (knowledge use) problem-solving processes;

- Students in Sub-group D2 applied the correct exploration strategy but failed in both the transition processes (score: 1 – 0 – 0);
- Students in Sub-group D3 also failed in both the transition processes but in a different way than those in Sub-group D2. Similarly to students in Group C, they failed in the exploration phase but managed to build the right concept map, that is, representing the problem as a result of guessing or misinterpreting their exploration. After building the right concept map, they were not able to transfer the guessing-based knowledge to the knowledge use part of the problem (score: 0 –1 –0);
- Students in Sub-group D4 also failed in both the transition processes but in a qualitatively different way than those in Sub-groups D2 and D3. They failed completely in the first transition process without demonstrating any understanding of the problem structure. However, after building the right concept map, they solved the application part of the problem most likely as a result of a trial-and-error strategy or very quick understanding but not as a result of any transition procedures from earlier phases of the problem-solving process (score: 0 – 0 – 1);
- Students in Sub-group D5 failed in the exploration phase and in both the transition processes, including all subsequent problem-solving processes (score: 0 – 0 – 0).

Fig. 3 synthesizes the transition-based approach and shows the transition-based groupings of the students on the problem level. This approach allows us to look behind the mechanisms of the students' learning and thinking processes during the problem-solving process on a very fine-grained level by providing more detailed pictures of the students' problem-solving behaviour via behavioural log data.

3. Results

3.1. The proportion of first-year university students who master both, only the first, only the second or neither of the two transitions in solving complex problems (RQ1a)

We analysed the relative frequencies of students in each problem scenario who belong to the four different groups of students based on their transition capabilities. Please note, because of the nested structure of the data (log data are nested in students) beyond the summative data describing the tendencies we provided problem-level results too. On average, one-third (31.9 %) of the university students mastered both transitions successfully (Group A, both transitions); however, the variance proved to be great between the different items. A total of 29.0 % of the students only had success with the first transition (Group B), and 0.2 % only managed the second transition (Group C). That is, there were almost no students who only mastered the second transition without mastering the first. 38.9 % of the students did not manage either of the transitions in their problem-solving (Group D – Sub-groups D1–D5). That is, mastering the first transition – correctly understanding and depicting the problem structure – played a crucial role in the problem-solving process, independently of any features (i.e., complexity, difficulty and position) of the problem. Most of the students who did not manage it failed in the following transition despite additional help (such as the presentation of the correct model in the knowledge application phase). The rate of successful transitions proved not to be stable across the problems (see RQ2), and it strongly varied by problem complexity and item position.

Table 1
The number of successful transitions by problem complexity.

Item/position/problem complexity*	Both transitions (Group A)	First transition only (Group B)	Second transition only (Group C)	No transitions (Group D)
Low	*63.5*	*14.5*	*0.9*	*21.1*
Item 1: 2 + 2+(2 + 0) [0, 0]	62.0	14.5	1.6	21.9
Item 2: 2 + 2+(2 + 0) [1, 0]	65.0	14.5	0.2	20.3
Medium	*36.8*	*43.0*	*0.0*	*20.2*
Item 3: 3 + 2+(3 + 0) [0, 0]	31.4	49.0	0.1	19.6
Item 4: 3 + 3+(3 + 0) [0, 0]	26.0	55.4	0.0	18.6
Item 5: 3 + 3+(4 + 0) [0, 0]	48.5	29.0	0.0	22.5
Item 6: 3 + 3+(4 + 0) [0, 0]	41.4	38.8	0.0	19.8
High	*11.1*	*22.2*	*0.0*	*66.7*
Item 7: 3 + 2+(2 + 1) [1, 0]	17.1	10.2	0.0	72.7
Item 8: 3 + 3+(3 + 1) [0, 1]	20.7	17.4	0.0	61.9
Item 9: 3 + 3+(3 + 1) [0, 1] (non-linear relation)	1.9	33.3	0.0	64.7
Item 10: 3 + 3+(3 + 1) [1, 0] (non-linear relation)	4.8	27.8	0.0	67.5

Note.
* Problem complexity includes number of input variables + number of output variables + (number of relations divided by number of direct relations + autonomous change) [number of irrelevant input variables + number of irrelevant output variables].

3.2. The number of successful transitions varies substantially with problem complexity (RQ1b)

The level of successful transitions proved to be quite stable on problems with similar complexity and varied substantially with problem complexity. Almost 64 % of the students mastered both transitions on problems of low complexity. This proportion dropped on problems of medium complexity (26–48 %), depending on item position, and dropped further on problems of high complexity. We have recognised major differences in transition capabilities amongst problems of high complexity. One-fifth of the students proved to be successful on highly complex problems with linear, autonomous changes, but just 2–5 % of the students were able to manage both of the transitions on problems with non-linear, autonomous changes. In contrast, almost 67 % of the students failed on both transitions on problems of high complexity, independently of the type of relations, one-fifth of them had no success on problems of medium complexity, and another one-fifth were unsuccessful on problems of low complexity (see Table 1). Thus, we can conclude that item complexity strongly influenced transitions throughout the problem-solving process, including all three processes, and we cannot expect that students behave the same way in each problem scenario, independently of its complexity and position.

3.3. Students' latent class transition profiles in complex problem-solving (RQ1c)

After theory-driven problem-level analyses of the different transition behaviours, we used the advantages of the nested structure of the data and monitored students real behaviour patterns on the test level. We used latent profile analyses to cluster students' transition behaviour as they worked with ten increasingly difficult complex problems. We assessed model specification for four latent class models. The information theory criteria used (AIC and BIC) indicated a continuous decrease in a growing number of latent classes. The likelihood ratio statistical test (Lo–Mendell–Rubin Adjusted Likelihood Ratio Test) showed the best model fit for the 4-class model. The entropy-based criterion reached the maximum values for the 3-class solutions, but it was also significant for the 4-class solutions. Based on the L–M–R test, we decided on the 4-class model to categorize 85 % of the first-year students into qualitatively different latent classes (Table 2) based on their transition capabilities while working on increasingly complex problems.

Fig. 4 shows the transition probabilities on the problem level and how they change during the ten, increasingly complex fictitious problems for students in different latent classes. Fig. 5 shows the rate of two, just one (the first) or no transitions problem by problem in each latent class. Based on Figs. 4 and 5, we can empirically distinguish students who are (1) expert transitioners, who manage to apply both of the transitions with the highest probability, almost independently of problem complexity; (2) advanced transitioners, who proved to be effective on low- and medium-complexity problems, but completely failed on problems of high complexity; (3) beginner transitioners, who generally managed to apply just the first transition but completely failed on the most complex problems; and (4) non-transitioners, who failed with regard to transitions with the highest probability on all of the problems, independently of problem complexity.

3.4. Characteristics of students' test-taking behaviour (i.e., time-on-task and number of interactions) vary according to the complexity and item position of the problem depending on the number of successful transitions (RQ2a)

For RQ1, we compared students' test-taking behaviour depending on their level and type of successful transition on each problem and built latent class profiles based on their transition behaviour pattern while working on ten complex problems. For RQ2, we examined how transition behaviour on both the problem (theory-driven approach) and test levels (latent transition profile-based approach) might additionally differ in the way students explore and attempt to solve complex problems.

As there were almost no students in Group C on any of the problems, we focused on Groups A, B and D on the problem level, that is, on the actual behaviour of students who mastered both transitions, only the first transition or no transitions on the problem level. The very first problem proved to be a "warm-up" task. Based on the ANOVA analyses, students' time-on-task behaviour showed the same pattern on all four problems of medium complexity, and they used almost the same amount of time, independently of their transition capabilities; however, there was a changing pattern that could be detected with regard to their interactions. The most successful students tended to engage in fewer interactions with more success than their peers after a learning process on the first two problems of medium complexity. The most complex problems changed students' transition behaviour pattern. Students in Groups A and B engaged in more interactions than those in Group D, independently of the type of relation (linear or non-linear), but they also used more time to do so on problems with non-linear, autonomous changes (see Appendix 2).

After closely examining Group D, whose members did not master any of the transitions successfully, we can distinguish between two statistically different test-taking behaviour profiles based on time-on-task and number of interactions, independently of problem

Table 2
Fit indices of the latent class models with differences in the specified number of latent profiles.

Index	Two	Three	Four	Five
Log likelihood	-13,566.64	-12,705.78	-12,484.42	-12,331.42
Free parameters	41	62	83	104
AIC	27,215.28	25,535.57	25,134.84	24,870.84
BIC	27,439.40	25,874.47	25,588.53	25,439.33
Entropy	.860	.903	.854	.826
LRT test (p)	4671.65 (0.00)	1710.80 (0.00)	439.00 (0.00)	304.05 (0.76)

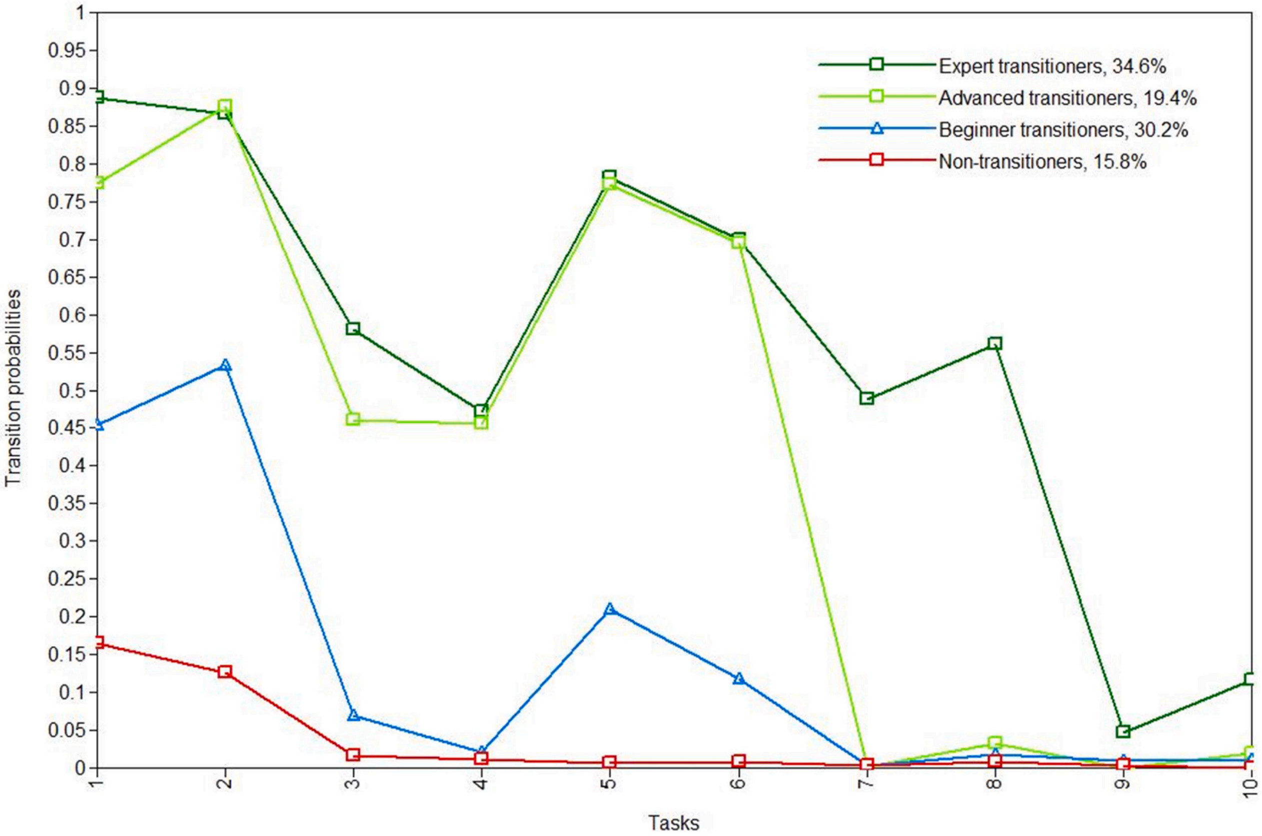

Fig. 4. Transition probabilities of expert, advanced, beginner and non-transitioners.

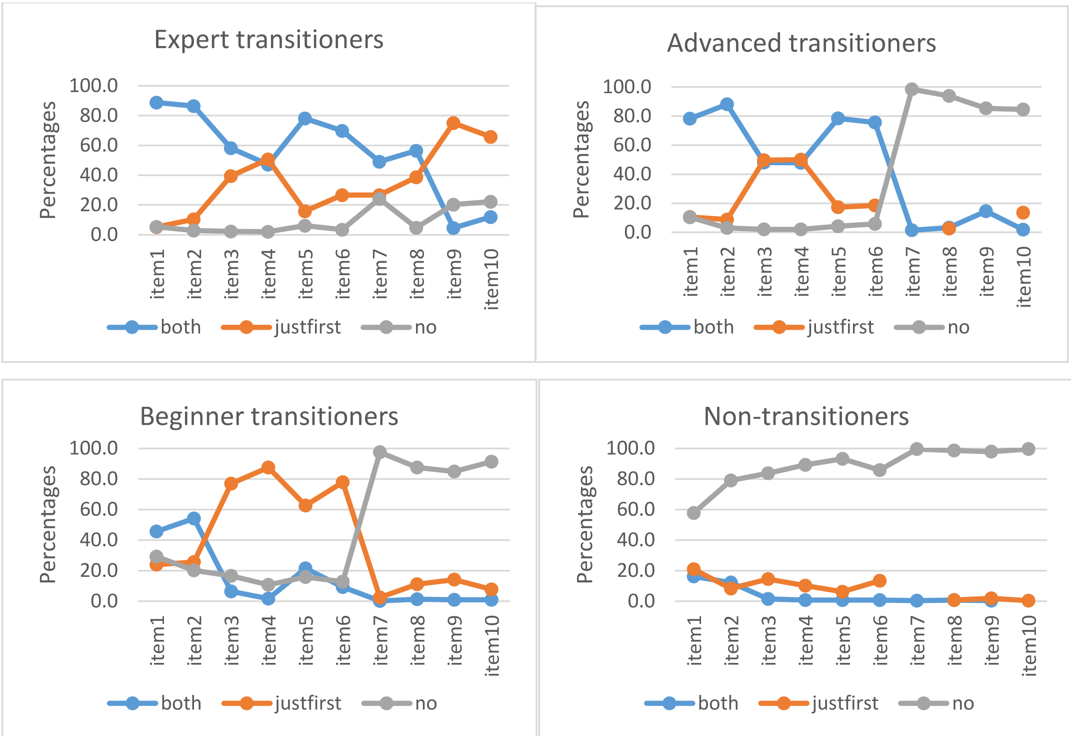

Fig. 5. The rate of both, just the first and neither of the transitions on the problem level in each latent class profile.

complexity (Appendix 3). Students who were effective in their problem exploration but failed in their interpretation and/or transition from problem exploration to problem representation (Sub-groups D1 and D2) behaved significantly differently than their peers who were not effective in their problem exploration (Sub-groups D3, D4 and D5). Students in the latter group – independently of problem complexity – spent significantly less time and clicked significantly less than students in D1 and D2 (see Appendix 3). That is, Group D is not a homogeneous group of test-takers. Students in the second set of sub-groups spent the least time and clicked the least throughout the problem-solving process, while students with an effective exploration strategy clicked more and spent more time solving problems than their peers in Groups A and B.

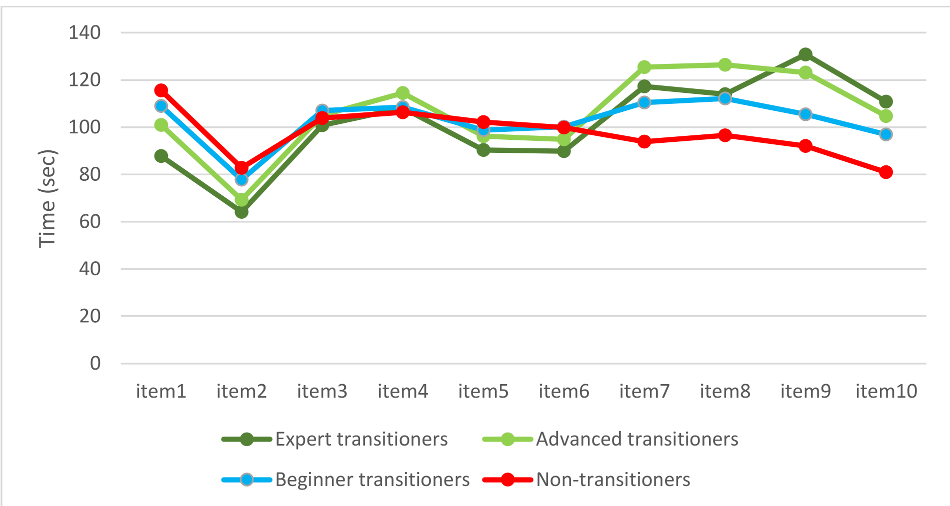

Fig. 6. The trends for time spent on the problem scenario by students with different latent class transition profiles.

3.5. Characteristics of students' test-taking behaviour (i.e., time-on-task and number of interactions) vary according to the latent class membership based on their transition capabilities (RQ2b)

Expert transitioners typically spent the shortest time solving the low- and medium-complexity problems, then they tended to use an increasing amount of time in parallel with the increasing complexity of the problems, and, finally, they devoted the longest time solving the last two, most complex problems. Advanced transitioners behaved like expert transitioners; that is, they typically spent less time on the less complex tasks and more time on the more complex ones. Apart from the two most complex problems, they spent more time on each problem scenario than their more successful peers. In contrast, members of the less successful groups typically spent the same amount of time on average on each problem scenario – after the warm-up task – independently of their complexity. This time proved to be the longest on the less complex problems and the shortest on the most complex ones compared to the trends detected in the three other latent transition class profiles. The behaviour pattern of the beginner transitioners was very similar to that of the non-transitioners on the low- and medium-complexity problems, but they tended to spend an increasing amount of time on the high-complexity items. To sum up, time spent on the problem scenario proved to be inversely proportional to the low-complexity problems and linearly proportional to the most complex problems with regard to latent class profile characteristics (Fig. 6).

The average number of interactions was very similar on the low- and medium-complex problems in each of the four latent classes but immensely different on the most complex problems. The number of interactions proved to be linearly proportional to the number of successful transitions (Fig. 7). That is, expert transitioners engaged in the most interactions, followed by the advanced, beginner and non-transitioners.

4. Discussion

The overarching aim of the study was to make knowledge transitions between the three problem-solving processes visible – i.e., the transition between (1) problem exploration and problem representation and that between (2) problem representation and knowledge use – and to determine the number and type of student's transition profiles based on their successful or unsuccessful transitions. In addition, we also monitored the characteristics of students' other test-taking behaviour process indicators, such as time-on-task and number of clicks, on students' transition skills as they solved problems in uncertain situations.

4.1. The magnitude and role of the transitions in solving complex problems of varying difficulty

The initial analysis showed that mastering the first transition – correctly understanding and depicting the problem structure – played a crucial role in the problem-solving process. Students who failed to master it regularly failed in the next transition. Results also indicated that problem complexity strongly influenced the transitions throughout the problem-solving process.

These results are consistent with previous research results that highlight the key role of the mental model transfer in knowledge

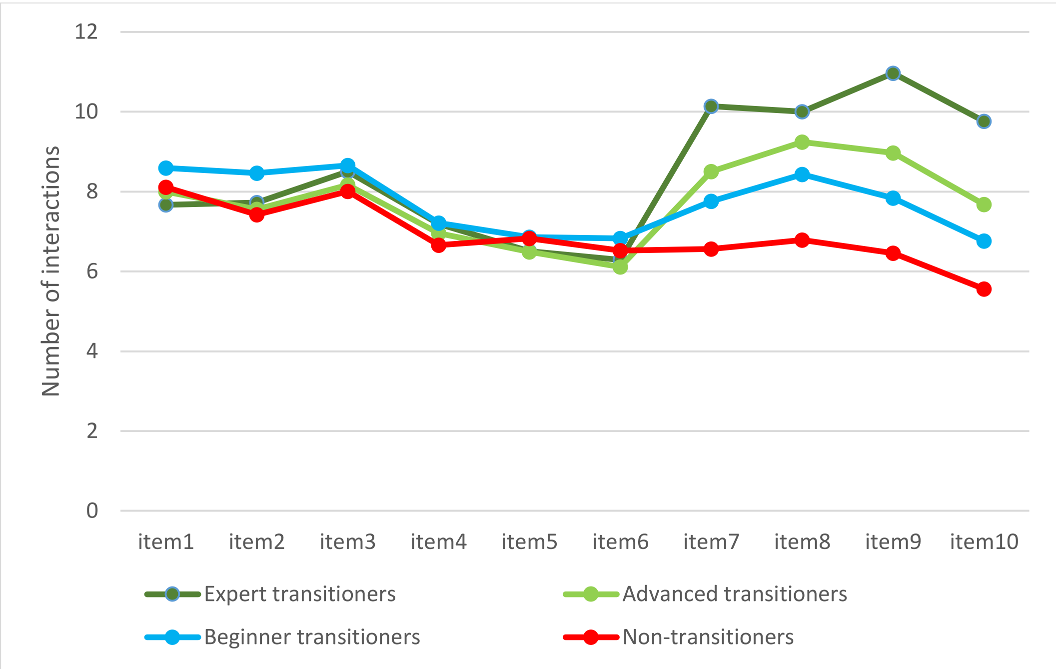

Fig. 7. The trends for interactions engaged in by students with different latent class transition profiles.

transition between the knowledge acquisition and knowledge application phases during problem-solving and the varying rates of successful knowledge transition and problem complexity (Nicolay et al., 2021). If we examine the transition processes more deeply, our results show that with the first knowledge transition, the effectiveness of the problem exploration and problem representation plays the most important role in problem-solving. That is, the correct understanding of the problem space and its representation in the form of the right mental model constitute one of the most vital steps in the problem-solving process, independently of problem complexity. Without a successful knowledge transition between problem exploration and problem representation – independently of problem complexity – it is not possible to make the second transition, the one between problem representation and knowledge use.

Using the nested structure of the data, based on students' real transition behaviour and behaviour pattern analyses, we have identified four latent transition class profiles, which describe students' transition behaviour more precisely and validly than an artificial grouping based on theory. As noted above, the four latent classes were: (1) expert transitioners, who mostly (not always) apply both transitions; (2) advanced transitioners, whose behaviour was similar to the effective transitioners on the low- and medium-complexity problems but failed on the most complex ones; (3) beginner transitioners, who mostly just used the first transition; and (4) non-transitioners.

4.2. The relation of the transitions to test-taking behaviour explains previous contradictory results on time-on-task, number of interactions and problem-solving achievement

First, we compared students' test-taking behaviour depending on their level and type of successful transition on each problem and on the test level. For RQ2, we examined how students in the different groups might additionally differ in the way they explore and attempt to solve complex problems. Specifically, we analysed features of students' test-taking behaviour (i.e., time-on-task and number of interactions) with respect to the number of successful transitions on the problem level and with respect to their complex behaviour patterns, that is, their latent transition class profile characteristics on the test level.

First, we focus on students' transition-level behaviour pattern and monitor their behavioural characteristics in terms of whether they manage to apply both, just the first or neither of the transitions. We concluded that students who managed to apply both transitions tended to engage in fewer interactions in less or the same amount of time with more success than students with less successful transitions in the same problem scenario. However, students' behaviour pattern changed on the most complex problems: to execute successful transitions, students needed more interactions and used more time than students who engaged in unsuccessful actions. The theory-based artificial Group D, which includes behavioural data for students who did not master any of the transitions successfully, proved to be a heterogeneous group based on their behavioural characteristics. Students who were effective in their problem exploration but failed in their transitions behaved significantly differently than their peers who were not effective in their problem exploration on a given problem. Students in the first set of sub-groups spent more time on problem-solving than their peers in the most successful group, while students in the second set of sub-groups spent the least time and clicked the least throughout the problem-solving process. That is, the number of interactions proved to be consistent with successful behaviour in their problem exploration, independently of their final achievement and transition skills. Students who applied a theoretically right exploration strategy in the problem exploration phase and succeeded or even failed in its interpretation typically engaged in more interactions than those who used a non-effective exploration strategy, independently of the problem environment. This indicates that number of interactions may be a good indicator of engagement and willingness to understand the problem space, but it provides less information about success in knowledge transition and problem-solving achievement.

The characteristics of the latent transition class profiles based on students' real behaviour patterns while solving all ten problems on the test validated and extended research results obtained on the transition and problem levels. Time spent on the problem scenario proved to be inversely proportional on the low-complexity problems and linearly proportional on the most complex problems; that is, effective transitioners typically spent the shortest time solving the low- and medium-complexity problems but the longest time solving the most complex problems. In contrast, members of the less successful groups typically spent the same amount of time, independently of their complexity. The average number of interactions proved to be a distinguishing characteristic on the most complex problems, where the number of interactions proved to be linearly proportional to the number of successful transitions.

These results partially explain previous contradictory research findings on time-on-task and number of interactions with problem-solving achievement and are consistent with Goldhammer et al. (2014) and Greiff et al.'s (2016) conclusions. Goldhammer et al. highlighted that even though time-on-task is a straightforward measure, the same amount of time may involve different behaviours. Greiff et al. (2016) argued that spending too much time on a task is associated with poor performance and suggested a U-shaped relation between time-on-task and problem-solving achievement.

Taking problem complexity into account, we concluded that students who managed to complete both transitions with the highest probability engaged in fewer interactions and spent less time on the low- and medium-complexity problems than their peers. They were fully aware of their interaction behaviour in their problem exploration and the meaning of the feedback provided by the system and did not need to use the trial-and-error strategy (Molnár & Csapó, 2018). This result was consistent using both theory-driven and profile-based approaches, as most of the students are expert transitioners, who managed to apply both transitions with high probability. In contrast, on the high-complexity problems, expert transitioners engaged in more interactions and spent more time in the problem-solving process, even if they failed at the very end. (Please note that expert transitioners had the highest probability of applying both transitions, but they did not manage to do so all the time in line with their profile characteristics.) They needed to use the trial-and-error strategy, which was time-consuming and called for more interactions.

The results of this study confirm that problem exploration and its successful interpretation and visualization in the form of a correct concept map play a crucial role in the transition between knowledge acquisition and knowledge application. For future problem-solving training programs, this means that supporting students' efforts in problem space exploration and spending time on the representation and visualization of their ideas about the structure of the problem space in the form of a mental model represent a key factor in developing problem-solving skills. This can be done in a number of ways, for example, by using the Kluge (2008) approach with only marginal guidance during problem space exploration and/or following Gopher et al., (1989) emphasis manipulation approach combined with Barrett et al. (2013) think-aloud protocol.

Another aspect of these results raises issues about the instrument used. To make the knowledge acquisition and the knowledge application phases independent, after the first phase, that is, after the processes of problem exploration and problem representation at the beginning of the application phase (process of knowledge use), students build the correct model, the correct problem representation in the form of a concept map. One can expect that this operation made the two transition processes independent. According to our results, the question can be raised whether it makes sense to actually present the correct model at the beginning of the application phase to the students because those who fail to master the first transition also fail to master the second one. They thus seem not to be able to make use of the model.

5. Limitations

The study used artificial, but, for measurement purposes, reliable and appropriate, problems developed with the MicroDYN approach; that is, results cannot be generalized to all kinds of everyday, complex and dynamic problems in real situations (see Funke, 2021). The study sample may lead to limitations and generalizability of the results on a population level. We used a non-representative convenience sample from one of the highest-ranked universities in Hungary. The study is basically a correlational study, so further research is needed on cause and effect. A further limitation of the results is that the data are nested in the students; that is, the observations in the different problem scenarios are not independent, but rather correlated on the student level.

6. Conclusion

The results of the current study provide important insights into how students transition between the different problem-solving processes – problem exploration, problem representation and knowledge use – and how this influences their overall problem-solving performance. For validation purposes, we also used learning analytics to investigate questions of the indicative power of test-taking behaviour on students' transition skills. As for the educational implications of developing problem-solving skills, supporting students' efforts in problem space exploration and spending time on visualizing their ideas about the problem structure, that is, supporting the first transition, the transition between problem exploration and problem representation, proved to be a key factor.

Author statement

No potential conflict of interest was reported by the authors. MG: Conceptualization, Methodology, Data Curation, Writing – Original draft preparation. SG: Writing-reviewing. SG is one of two authors of the commercially available COMPRO test that is based on the multiple complex systems approach and that employs the same assessment principle as MicroDYN. However, a free version of MicroDYN is available for any research and educational purposes. SG receives royalty fees for COMPRO.

Ethical statement

Ethical approval was not required for this study in accordance with national and institutional guidelines. The assessments which provided data for this study were integrated parts of the educational processes of the participating university. The participation was voluntary. All of the students in the assessment had turned 18 by the time of the assessment; that is, it was not required or possible to request and obtain written informed parental consent from the participants.

Funding

This research was supported by a Hungarian National Research, Development and Innovation Fund grant (under the OTKA K135727 funding scheme) and a Hungarian Academy of Sciences Research Programme for Public Education Development grant (KOZOKT2021-16). It was also funded by a grant from the Fonds National de la Recherche Luxembourg (Luxembourg National Research Fund) [CORE 'TRIOPS'].

Data availability

Data will be made available on request.

Aknowledgements

Beno Csapo, our esteemed colleague and mentor, deceased while we were working on this paper. We mourn the loss of a wonderful person and a great researcher. This paper is dedicated to his legacy.

Appendix 1. Examples of task descriptions and linear structural equations for low-, medium- and high-complexity problems

Complexity	Input	Output	Relations	Eigen dynamic	Linear structural equations
Low	2	2	2	no	$X_t + 1 = 1 \times X_t + 2 \times A_t + 0 \times B_t$ $Y_t + 1 = 1 \times Y_t + 0 \times A_t + 2 \times B_t$
Medium	3	3	4	no	$X_t + 1 = 1 \times X_t + 2 \times A_t + 0 \times B_t + 0 \times C_t$ $Y_t + 1 = 1 \times Y_t + 0 \times A_t + 2 \times B_t + 2 \times C_t$ $Z_t + 1 = 1 \times Z_t + 0 \times A_t + 2 \times B_t + 0 \times C_t$
High	3	3	4	yes	$X_t + 1 = 1 \times X_t + 0 \times A_t + 2 \times B_t + 0 \times C_t$ $Y_t + 1 = (1 \times Y_t + 2 \times A_t + 0 \times B_t + 0 \times C_t) + 3$ $Z_t + 1 = 1 \times Z_t + 0 \times A_t + 0 \times B_t + 2 \times C_t$

Appendix 2

	Group A		Group B		Group D		ANOVA
Time	Mean	SD	Mean	SD	Mean	SD	
Item 1: 2 + 2+(2 + 0) [0, 0]	91.53	37.68	114.28	44.75	120.93	52.57	{A} < {B, D}
Item 2: 2 + 2+(2 + 0) [1, 0]	66.16	29.17	74.43	34.08	90.40	47.83	{A, B, D}
Item 3: 3 + 2+(3 + 0) [0, 0]	97.50	35.27	106.48	38.00	108.71	53.17	{A, B, D}
Item 4: 3 + 3+(3 + 0) [0, 0]	107.64	27.34	110.58	41.05	106.96	50.62	{A, B, D}
Item 5: 3 + 3+(4 + 0) [0, 0]	90.25	31.89	98.09	36.32	105.36	50.46	{A, B, D}
Item 6: 3 + 3+(4 + 0) [0, 0]	88.21	25.09	100.97	36.65	100.19	46.54	{A} < {B, D}
Item 7: 3 + 2+(2 + 1) [1, 0]	111.94	32.99	121.06	43.35	112.17	47.57	{A, D} < {B}
Item 8: 3 + 3+(3 + 1) [0, 1]	112.91	32.03	114.31	38.27	112.73	45.80	{A, B, D}
Item 9: 3 + 3+(3 + 1) [0, 1] (non-linear relation)	121.38	28.06	126.79	39.15	109.42	45.47	{D} < {A, B}
Item 10: 3 + 3+(3 + 1) [1, 0] (non-linear relation)	111.59	36.91	110.50	36.80	95.84	40.92	{D} < {A, B}
Complexity low	78.54	35.90	94.36	44.46	106.24	52.57	{A} < {B} < {D}
Complexity medium	94.29	30.94	105.15	38.87	105.27	50.29	{A, B, D}
Complexity high	112.77	32.79	118.58	39.38	107.50	45.55	{D} < {A} < {B}
Interactions_total							
Item 1: 2 + 2+(2 + 0) [0, 0]	7.70	3.64	9.61	4.73	8.51	5.37	{A, B, D}
Item 2: 2 + 2+(2 + 0) [1, 0]	7.54	3.77	8.68	4.50	8.41	6.08	{A, B, D}
Item 3: 3 + 2+(3 + 0) [0, 0]	8.37	3.58	8.51	3.48	8.23	4.71	{A, B, D}
Item 4: 3 + 3+(3 + 0) [0, 0]	7.02	1.88	7.20	3.02	6.76	3.44	{D} < {A, B}
Item 5: 3 + 3+(4 + 0) [0, 0]	6.37	2.55	6.84	3.07	7.07	4.14	{A} < {B, D}
Item 6: 3 + 3+(4 + 0) [0, 0]	6.13	1.89	6.79	2.77	6.50	3.24	{A} < {B, D}
Item 7: 3 + 2+(2 + 1) [1, 0]	10.02	3.12	10.92	4.41	7.85	3.73	{D} < {A, B}
Item 8: 3 + 3+(3 + 1) [0, 1]	9.99	3.54	9.76	3.89	8.24	3.84	{D} < {A, B}
Item 9: 3 + 3+(3 + 1) [0, 1] (non-linear relation)	10.26	3.85	10.83	3.63	7.89	3.70	{D} < {A, B}
Item 10: 3 + 3+(3 + 1) [1, 0] (non-linear relation)	10.23	3.66	10.00	3.61	6.69	3.41	{D} < {A, B}
Complexity low	7.62	3.71	9.15	4.63	8.46	5.72	{A, D} < {D, B}
Complexity medium	6.84	2.68	7.42	3.19	7.14	3.99	{A, B, D}
Complexity high	10.04	3.41	10.37	3.80	7.66	3.25	{D} < {A} < {B}

Appendix 3

	A	B	D1	D2	D3	D4	D5	ANOVA
Time_total	90.59	108.15	113.72	111.57	75.37	81.75	75.35	{D3, D5} < {A, D4}<{B, D1, D2}
Item 1: 2 + 2 + 2	91.53	114.28	119.40	137.00	78.20	89.71	94.46	{A, D3, D4, D5} < {B, D1, D2}
Item 2: 2 + 2 + 2	66.16	74.43	97.96	96.94	–	59.57	65.79	{A, B, D1, D2, D4, D5}
Item 3: 3 + 2 + 3	97.50	106.48	122.65	119.95	69.73	–	81.39	{A, D3, D5} < {B, D1, D2}
Item 4: 3 + 3 + 3	107.64	110.58	151.00	118.89	73.12	79.84	79.84	{D3, D5} < {A, D4} < {B, D1, D2}
Item 5: 3 + 3 + 4	90.25	98.09	114.09	111.13	–	93.50	79.04	{A, D4, D5} < {B, D1, D2}
Item 6: 3 + 3 + 4	88.21	100.97	106.22	107.99	–	68.25	82.22	{A, D4, D5} < {B, D1, D2}
Item 7: 3 + 2+(2 + 1)	111.94	121.06	118.06	114.07	–	99.56	74.65	{D5} < {A, B, D1, D2, D4}
Item 8: 3 + 3+(3 + 1)	112.91	114.31	122.01	115.05	–	–	74.88	{D5} < {A, B, D1, D2}
Item 9: 3 + 3+(3 + 1)	121.38	126.79	110.40	112.93	–	–	73.69	{A, B, D1, D2, D5}
Item 10: 3 + 3+(3 + 1)	111.59	110.50	98.74	99.94	–	86.50	61.90	{D4, D5} < {A, B, D1, D2}
Complexity low	78.54	94.36	110.07	117.05	76.05	68.18	77.66	{A, D3, D4, D5} < {B, D1, D2}
Complexity medium	94.29	105.15	114.92	115.32	71.47	86.30	80.60	{A, D3, D4, D5} < {B, D1, D2}
Complexity high	112.77	118.58	115.64	110.30	91.22	97.67	70.57	{D3, D4, D5}<{A, D1, D2} < {B}
Interaction_total								
Item 1: 2 + 2 + 2	7.70	9.61	8.92	9.83	3.27	4.29	3.75	{D3, D4, D5} < {A, B, D1, D2}

(*continued on next page*)

(*continued*)

	A	B	D1	D2	D3	D4	D5	ANOVA
Item 2: 2 + 2 + 2	7.58	8.92	8.38	9.04	–	5.29	5.21	{D3, D4, D5} < {A, B, D1, D2}
Item 3: 3 + 2 + 3	8.30	9.39	8.39	8.89	4.93	–	5.62	{D3, D5} < {A, B, D1, D2}
Item 4: 3 + 3 + 3	7.03	7.40	7.36	7.60	3.40	4.64	4.33	{D3, D5} < {A, B, D1, D2, D4}
Item 5: 3 + 3 + 4	6.56	7.12	7.04	7.06	–	5.07	5.04	{D5} < {A, B, D1, D2, D4}
Item 6: 3 + 3 + 4	6.35	7.12	6.63	6.70	–	5.43	4.75	{D5} < {A, B, D1, D2, D4}
Item 7: 3 + 2+(2 + 1)	9.10	8.40	7.60	7.19	–	6.07	5.46	{D5} < {A, B, D1, D2, D4}
Item 8: 3 + 3+(3 + 1)	9.35	8.61	8.39	8.00	–	–	5.00	{D5} < {A, B, D1, D2}
Item 9: 3 + 3+(3 + 1)	9.67	8.05	8.09	7.76	–	–	4.50	{D5} < {A, B, D1, D2}
Item 10: 3 + 3+(3 + 1)	8.51	7.00	6.83	6.60	–	3.71	4.67	{D5} < {A, B, D1, D2, D4}
Complexity low	7.62	3.71	9.32	9.68	3.05	3.45	3.81	{B, D3, D4, D5} < {A, D1, D2}
Complexity medium	6.84	2.68	7.40	8.53	4.25	4.19	3.99	{B, D3, D4, D5} < {A, D1, D2}
Complexity high	10.04	3.41	8.49	7.98	5.78	5.35	3.67	{B, D3, D4, D5} < {A, D1, D2}

References

Alzoubi, O., Fossati, D., Di Eugenio, B., Green, N., & Chen, L.. Predicting students' performance and problem solving behavior from iList log data. https://scholar.google.com/scholar_lookup?title=Predicting%20students%20performance%20and%20problem%20solving%20behavior%20from%20iList%20log%20data&publication_year=2013&author=O.%20Alzoubi&author=D.%20Fossati&author=B.%20Di%20Eugenio&author=N.%20Green&author=L.%20Chen.

Barrett, J. D., Peterson, D. R., Hester, K. S., Robledo, I. C., Day, E. A., Hougen, D. P., & Mumford, M. D. (2013). Thinking about applications: Effects on mental models and creative problem-solving. *Creativity Research Journal, 25*(2), 199–212. https://doi.org/10.1080/10400419.2013.783758

Blair, C., & Razza, P. R. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*, 647–663. 10.1111/j.1467-8624.2007.01019.x.

Chan, J. Y. C., Ottmar, E. R., & Lee, J. E. (2022). Slow down to speed up: Longer pause time before solving problems relates to higher strategy efficiency. *Learning and Individual Differences, 93*, Article 102109.

Csapó, B., & Molnár, G. (2019). Online Diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in Psychology*.

Csapó, B., Lőrincz, A., & Molnár, G. (2012). Innovative assessment technologies in educational games designed for young students. *Assessment in game-based learning* (pp. 235–254). New York, NY: Springer.

Eichmann, B., Goldhammer, F., Greiff, S., Brandhuber, L., & Naumann, J. (2020a). Using process data to explain group differences in complex problem solving. *Journal of educational psychology, 112*(8), 1546.

Eichmann, B., Greiff, S., Naumann, J., Brandhuber, L., & Goldhammer, F. (2020b). Exploring behavioural patterns during complex problem-solving. *Journal of Computer Assisted Learning, 36*(6), 933–956.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving, 4*, 19–42. 10.7771/1932-6246.1118.

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning, 7*(1), 69–89. 10.1080/13546780042000046.

Funke, J. (2021). It requires more than intelligence to solve consequential world problems. *Journal of Intelligence, 9*(3), 38.

Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology, 106*(3), 608–626. https://doi.org/10.1037/a0034716

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica, 71*(1–3), 147–177. https://doi.org/10.1016/0001-6918(89)90007-3

Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education, 126*, 248–263.

Greiff, S., Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving. An analysis of behavioral data from computer-generated log files. *Computers in Human Behavior, 61*, 36–46. 10.1016/j.chb.2016.02.095.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new assessment perspective. *Applied Psychological Measurement, 36*(3), 189–213.

Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B (2013). Complex problem solving in educational contexts – Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology, 105*(2), 364–379. 10.1037/a0031856.

Heinonen, J., Aro, T., Ahonen, T., & Poikkeus, A.-M. (2011). Test-taking behaviors in a neurocognitive assessment: Associations with school-age outcomes in a Finnish longitudinal follow-up. *Psychological Assessment, 23*(1), 184–192. https://doi.org/10.1037/a0021291

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*, 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Kluge, A. (2008). What you train is what you get? Task requirements and training methods in complex problem-solving. *Computers in Human Behavior, 24*(2), 284–308. https://doi.org/10.1016/j.chb.2007.01.013

Lotz, C., Scherer, R., Greiff, S., & Sparfeldt, J. R. (2017). Intelligence in action–Effective strategic behaviors while solving complex problems. *Intelligence, 64*, 98–112. https://doi.org/10.1016/j.intell.2017.08.002

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., … Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128–165. 10.1037/0003-066X.56.2.128.

Molnár, G., & Csapó, B. (2017). Exploration and learning strategies in an interactive problem-solving environment at the beginning of higher education studies. In J. C. Spender, G. Schiuma, & T. Gavrilova (Eds.), *Knowledge Management in the 21st Century: Resilience, Creativity and Co-creation. Proceedings of 12th edition of the International Forum on Knowledge Asset Dynamics* (pp. 283–292). St. Petersburg: St. Petersburg University.. http://publicatio.bibl.u-szeged.hu/11591/1/Molnar_Csapo_2017_ExplorationAndLearningStrategies.pdf

Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: logfile analyses. *Frontiers in Psychology, 9*, 302.

Molnár, G. (2021). How to make different thinking profiles visible through technology: The potential for log file analysis and learning analytics. In M. Virvou, G. A. Tsihrintzis, L. H. Tsoukalas, & L. C. Jain (Eds.), *Advances in artificial intelligence-based technologies* (pp. 125–146). Cham: Springer.

Molnár, G., Alrababah, S. A., & Greiff, S. (2022). How we explore, interpret, and solve complex problems: A cross-national study of problem-solving processes. *Heliyon, 8*, e08775.

Mustafić, M., Yu, J., Stadler, M., Vainikainen, M. P., Bornstein, M. H., Putnick, D. L., & Greiff, S. (2019). Complex problem solving: Profiles and developmental paths revealed via latent transition analysis. *Developmental psychology, 55*(10), 2090.

Nasir, J., Kothiyal, A., Bruno, B., & Dillenbourg, P. (2021). Many are the ways to learn identifying multi-modal behavioral profiles of collaborative learning in constructivist activities. *International Journal of Computer-Supported Collaborative Learning, 16*(4), 485–523. 10.1007/s11412-021-09358-2.

Naumann, J., & Goldhammer, F. (2017). Time-on-task effects in digital reading are non-linear and moderated by persons' skills and tasks' demands. *Learning and Individual Differences, 53*, 1–16. https://doi.org/10.1016/j.lindif.2016.10.002

Newton, K. J., Lange, K., & Booth, J. L. (2020). Mathematical flexibility: Aspects of a continuum and the role of prior knowledge. *The Journal of Experimental Education, 88*(4), 503–515. https://doi.org/10.1080/00220973.2019.1586629

Nicolay, B., Krieger, F., Stadler, M., Gobert, J., & Greiff, S. (2021). Lost in transition – Learning analytics on the transfer from knowledge acquisition to knowledge application in complex problem solving. *Computers in Human Behavior, 115*. https://doi.org/10.1016/j.chb.2020.106594

Schweizer, F., Wüstenberg, S., & Greiff, S. (2013). Validity of the MicroDYN approach: Complex problem solving predicts school grades beyond working memory capacity. *Learning and Individual Differences, 24*, 42–52.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review, 84*(2), 127–190. https://doi.org/10.1037/0033-295X.84.2.127

Stadler, M., Hofer, S., & Greiff, S. (2020). First among equals: Log data indicates ability differences despite equal scores. *Computers in Human Behavior, 111*. https://doi.org/10.1016/j.chb.2020.106442

Stadler, M., Niepel, C., & Greiff, S. (2016). Easily too difficult: Estimating item difficulty in computer simulated microworlds. *Computers in Human Behavior, 65*, 100. https://doi.org/10.1016/j.chb.2016.08.025

Wu, H., & Molnár, G. (2021). Logfile analyses of successful and unsuccessful strategy use in complex problem-solving: A cross-national comparison study. *European Journal of Psychology of Education*.

Zhang, X., Meng, Y., de Pablos, P. O., & Sun, Y. (2019). Learning analytics in collaborative learning supported by Slack: From the perspective of engagement. *Computers in Human Behavior, 92*, 625–633.