

A Pilot Study for Measuring Complex Problem-Solving in Jordan: Feasibility, Construct Validity, and Behavior Pattern Analyses

SAGE Open
April-June 2024: 1–13
© The Author(s) 2024
DOI: 10.1177/21582440241249884
journals.sagepub.com/home/sgo
 Sage

Saleh Ahmad Alrababah¹ , Hao Wu^{2,3} , and Gyöngyvér Molnár^{3,4}

Abstract

In the 21st century, complex problem-solving (CPS) serves as a key indicator of educational achievement. However, the elements of successful CPS have not yet been fully explored. This study investigates the role of strategic exploration and different problem-solving and test-taking behaviors in CPS success, using logfile data to visualize and quantify students' problem-solving behavior on 10 CPS problems with different characteristics and levels of difficulty. Additionally, in the present study, we go beyond the limits of most studies that focus on students' problem-solving behavior pattern analyses in European cultures and education systems to examine Arabic students' CPS behavior. The results show that computer-based assessments of CPS are feasible and valid in Jordanian higher education. The findings also confirm the structural validity of CPS, indicating that the processes of knowledge acquisition (KAC) and knowledge application (KAP) can be distinguished and separated in the problem-solving process. Large differences were identified in students' test-taking behavior in terms of the efficacy of their exploration strategy. We identified four latent classes based on the students' exploration strategy behavior. The study thus leads to a better understanding of how students solve problems and behave during the problem-solving process in uncertain situations.

Keywords

complex problem-solving, logfile analysis, test-taking behavior, higher education, exploration strategy

Introduction

Problem-solving is part of our daily lives. For example, we have to solve the following problems every day: what to wear, how to get to our destination (e.g., work or school), how to book tickets for a train, for a bus, for a journey, where and what to eat for lunch, and so on. We constantly face problems that require the application of problem-solving skills. Some of these problems can be solved very easily based on our earlier experiences, while others are more complex, requiring new information and more complex decision-making processes. In the past, education focused on teaching and assessing factual knowledge almost exclusively. It typically favored memorization as a key learning strategy. Nowadays, however, more emphasis is placed on knowledge creation and application, as students have their own devices hooked up to the Internet for easy access to information.

Society and the environment are continually changing (Griffin et al., 2012), and technologies are rapidly

evolving in almost every area of business, thus resulting in content knowledge quickly becoming obsolete. According to the OECD (2014a), “[a]dapting, learning, daring to try out new things and always being ready to learn from mistakes are among the keys to resilience and success in an unpredictable world” (p. 13). Thus, problem-solving has become one of the most essential 21st-century skills (Dede, 2010), whose development should be one of the primary goals of education

¹Doctoral School of Education, University of Szeged, Hungary

²School of Education, University of Saint Joseph, Macao

³Institute of Education, MTA-SZTE Digital Learning Technologies Research Group, University of Szeged, Hungary

⁴SZTE IKIKK Digital Learning Technologies Incubation Research Group, Hungary

Corresponding Author:

S. Alrababah, Doctoral School of Education, University of Szeged, Petőfi Sándor sgt. 32–34, Szeged H-6722, Hungary.
Email: salehrabah5@gmail.com



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of

the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

(Molnár et al., 2022). Schools should prepare their students for jobs and technologies that do not yet exist and solve problems that have never been faced before (OECD, 2018) to succeed in this new world. Those prospects represent novel needs in higher education and have led to a growing interest in assessment instruments that cover a broader area of competencies than traditional domain-specific skills and disciplinary knowledge (Molnár & Csapó, 2018). These assessment instruments can be used to measure students' 21st-century skills (i.e., collaboration skills as a measure of teamwork, decision-making, creativity, information literacy, critical thinking, and problem-solving; van Laar et al., 2020).

The present study focuses on the assessment of problem-solving, especially complex problem-solving (CPS), in an Arabic higher education environment. Through the assessment of problem-solving skills, we aim to monitor the suitability of education programs in terms of the development of students' 21st-century skills in Jordan, thus obtaining more knowledge about the factors and mechanisms that constitute a successful complex problem-solver, while considering the problem-solving behavior of students socialized in different cultures. This study therefore investigates different factors, such as the role of strategic exploration and test-taking behavior in terms of successful strategy usage in a CPS environment, using logfile data to visualize and quantify students' problem-solving behavior on 10 CPS problems with different levels of difficulty and characteristics. Our paper is among the first (see e.g., Molnár et al., 2022) to study the feasibility and validity of an interactive, innovative, third-generation computer-based test, such as the globally examined CPS assessments (see e.g., Csapó & Funke, 2017; Csapó et al., 2014; Dörner & Funke, 2017; Molnár & Csapó, 2018; Molnár et al., 2017; OECD, 2014a; Wu and Molnár, 2018; Wu & Molnár, 2021; Wüstenberg et al., 2014) in Jordanian higher education. Despite the great attention paid to CPS internationally, there is a lack of it in countries in the Arab region (Mousa and Molnár, 2019; Mousa and Molnár, 2020), especially in Jordan, where computer-based assessment (CBA) has less of a tradition compared to other countries (Molnár et al., 2022). We investigated whether Jordanian students interpret problems the same way as students in other, mostly European, countries (Greiff et al., 2013, 2015; OECD, 2014a; Wüstenberg et al., 2014), where most CPS studies have been carried out (Molnár et al., 2022). Beyond students' CPS performance, CBA makes it possible to monitor additional test-taking behavioral actions, such as mouse clicks, time-on-task, and problem-solving strategy (Gnaldi et al., 2020). These sorts of information have the potential to

provide policymakers and researchers with valuable insights into students' CPS skills and offer new ways to assist them in optimizing their cognitive capacity (Wu & Molnár, 2021). Such information is still missing from the evaluation and development of different education systems with increasing cultural diversity.

Theoretical Background

Complex Problem-Solving and its Assessment

The classical view defines problem-solving as a step-by-step process, which is passive, reproductive, and domain-general, mostly based on trial and error (Greiff et al., 2013). In contrast, the Gestalt view considers problem-solving as a productive and active process, where insight, reorganization, and functional fixedness play an important role (Schnitz et al., 2010). The development of the information-processing approach and Newell and Simon's problem space theory have opened the door to new directions in research. North American research has typically focused on examining the development of expertise in separate domains, while most of the research in Europe has concentrated on the problem-solving processes of complex, unknown problems through computerized scenarios. Reiff et al. (2006) defined problem-solving as guided thinking and action in situations with no routine solution. Eichmann et al. (2020) distinguished between analytical and interactive problem-solving according to the interactive nature of the problem scenario.

In analytical (static) problem-solving environments, both the problem and the related information are static. That is, there are no changes during the problem-solving process, with all the relevant information being presented at the beginning of the process (Greiff et al., 2013).

CPS requires a sequence of complex cognitive processes which a person employs to derive new information from the problem context with the intention of making decisions, solving problems, and implementing plans of action. This requires continuous activities (Funke, 2010) during the problem-solving process. In the history of CPS assessments, we can distinguish between two different approaches to measuring CPS (Buchner, 1995; Funke, 2014):

- (1) Computer-simulated microworlds, which have a large number of variables like real-life problems. For example, the well-known microworld scenario "Lohhausen," which consists of nearly 2,000 associated variables (Dörner 1983). This approach results in highly complex problems with high-level similarities to real-world problems. However, (a) their application requires a very long testing time, and (b) they fail to employ common theoretical frameworks to

produce comparable problems in a systematic way (Funke, 2001; Funke & Frensch, 2007). In addition, (c) participants' performance is influenced by many other factors, such as prior knowledge about the problem context, not only their problem-solving skills (Greiff et al., 2015). Finally, (d) the majority of microworld-based problems consist of a few items or many interconnected items, both harming instrument reliability (Greiff et al., 2015).

- (2) Simplified and artificial but still complex problems that follow specific rules of construction. Most (though not all) of the characteristics of a complex system are present in minimal complex systems (dynamic, complex, and intransparent; see Funke, 1991). A minimal complex system has a low number of variables and relations, resulting in reduced testing time compared to the highly complex and challenging microworlds. The MicroDYN approach falls into this category (Greiff & Funke, 2009; Greiff et al., 2012; Schweizer et al., 2013). It employs a number of independent "fake" scenarios to prevent the influence of participants' previous knowledge (Greiff et al., 2015), it uses only a few variables—that is, problems are easy to scale—and it is widely accepted among problem-solving assessments (see e.g., Csapó & Molnár, 2017; Greiff & Wüstenberg, 2014; Greiff et al., 2015; Mustafić et al., 2019; OECD, 2014b).). However, there are limitations to generalization to consider as regards problems of minimal complexity in comparing real-life problems because variables cannot be selectively controlled in a real-life context in most cases (Funke, 2021).

The focus of the present study is on CPS, especially the MicroDYN approach (Funke, 2014), measured in a computerized and interactive environment. The MicroDYN approach was proposed by Joachim Funke and operationalized by Samuel Greiff in the form of a CBA tool (see e.g., Funke, 2001; Greiff & Funke, 2009). According to the theoretical understanding, CPS in the MicroDYN approach is a two-dimensional construct (Funke, 2001; Greiff et al., 2012, 2013; Leutner et al., 2005), consisting of knowledge acquisition (KAC) and knowledge application (KAP). In the first phase of the problem-solving process, the problem-solver needs to acquire knowledge in uncertain situations (KAC), while, in the second phase, this newly acquired knowledge must be applied in a goal-directed way toward the problem solution (Funke, 2001; Greiff et al., 2018; Novick & Bassok, 2005). In a real-life setting, these two processes are related and take place at the same time. However, in an assessment situation, they are usually separated.

The Role of Strategic Exploration in Problem-Solving

Studies on CPS have focused on describing the different strategies people use to explore CPS problem environments. Van Der Linden et al. (2001) determined five effective exploration strategies: (1) the optimal strategy was a systematic exploration, while (2) trial-and-error, (3) rigid exploration, (4) encapsulation, and (5) negative self-evaluation were considered sub-optimal strategies. The role and success of systematic exploration were confirmed by Molnár and Csapó (2018), who highlighted that the application of systematic exploration led to the solution with the highest probability. The term *systematic* indicates the process by which a person sets goals based on assumptions about how a task should be accomplished. The trial-and-error strategy was expected to be less effective (Dörner & Wearing, 1995). The key to this strategy is a lack of assumptions that direct behavior. Rigid exploration is one behavior that is frequently observed in problem-solving (Hollnagel, 1993). The action sequences in this type of behavior are typically repetitive. Even though there is strong evidence that certain assumptions are wrong, people who work rigidly tend to hold onto them for a long time. Encapsulation in information seeking was another sort of behavior that can be observed when attempting to solve a complex problem. The drawback of this strategy is that collecting information becomes the major goal. Low performance and sub-optimal strategies have also been associated with negative self-evaluation strategy, since fewer attentional resources would be available to work on the main task when people focus on negative cognitions (Mikulincer, 1989).

Exploring and generating effective information represent the secret to solving a problem successfully. According to Wittmann and Hatrup (2004), "riskier strategies [create] a learning environment with greater opportunities to discover and master the rules and boundaries [of the problem]" (p. 406). Thus, there may be differences in the efficacy of the exploration strategies when gathering information about a problem (Wu & Molnár, 2021). Problem-solvers are supposed to explore the problem environment by acquiring knowledge during strategic exploration (Fischer et al., 2012). The development and implementation of strategic exploration are central actions of the problem-solving process (Wüstenberg et al., 2012). Exploring and creating valuable information are essential to solving problems successfully (Wüstenberg et al., 2012).

Problem-solving success in MicroDYN scenarios, which are simplifications and simulations of real-world problems, is also affected by the adoption and application of strategic exploration. In these artificial problem situations, the isolated variation strategy has been the most frequently discussed exploration strategy (it is often called the vary-one-thing-at-a-time (VOTAT) strategy;

Vollmeyer et al., 1996). Using the VOTAT strategy, the problem-solver directly detects the effects of a single variable at a time by manipulating that variable in a systematic way while keeping the other variables unchanged, that is, in the neutral position (Molnár & Csapó, 2018). According to previous studies, participants who know how to apply VOTAT are more likely to achieve better on problem-solving tasks (Greiff et al., 2018), particularly in minimal complex systems (Fischer et al., 2012). According to Lotz et al. (2017), effective use of VOTAT correlates with higher levels of intelligence and successful exploration behavior may lead to better results in problem-solving (Wu & Molnár, 2021).

Molnár and Csapó (2018) conducted an empirical study to investigate how students' exploration strategies affect their performance in a CPS environment. They assessed the problem-solving achievement of a group of 3rd- to 12th-grade (ages 9–18) Hungarian students ($N = 4,371$) and modeled the participants' exploration strategies. This finding supported the notion that students' problem-solving performance is influenced by their exploration strategies. For example, conscious VOTAT strategy users achieved better on a CPS test than their peers. Additionally, other empirical studies (e.g., Molnár et al., 2022; Wu & Molnár, 2021) produced similar results, emphasizing the significance of VOTAT in a MicroDYN-based CPS environment.

VOTAT is among the most effective exploration strategies in most problem-solving environments (Lotz et al., 2017; Wu & Molnár, 2021), and it is the most effective in minimal complex systems (such as MicroDYN). Based on Greiff et al. (2018) and Molnár and Csapó (2018), we have discerned and quantified three types of exploration strategies in each of the problem scenarios in the present analyses: (1) no VOTAT (no VOTAT trial was applied); (2) partial VOTAT (VOTAT trials were used for some but not all of the variables in a given problem scenario); and (3) full VOTAT (VOTAT trials were applied for all of the variables in a given CPS scenario) (see Greiff et al., 2018; Molnár & Csapó, 2018; Wu & Molnár, 2021).

Research Questions

Nowadays, there is a positive attitude toward using technology in higher education in Jordan (Al-Khayat, 2017), but we do not have any proof of its feasibility and applicability, especially in the area of assessment. Thus, at the initial phase of the study, we had to test the feasibility and applicability of using innovative, interactive, third-generation computer-based tests in Jordan in an educational context. We also explored students' test-taking and problem-solving behavior while solving complex problems in a digital environment with both directly collected answer data and logfile analyses. Toward our

objectives of determining whether we had structural validity and whether CPS was a two-dimensional construct in the Jordanian educational context as well as to ascertain the characteristics of the mechanism underlying successful problem-solving, among other elements, we aimed to answer the following research questions:

Research question 1: Is an interactive, innovative, third-generation CPS test reliable and applicable in a country where CBA has a relatively short history?

Research question 2: To what degree do the CPS scores reflect the dimensionality of the construct being measured in the Jordanian educational context?

Research question 3: What types of strategic exploration strategies were used by the Jordanian university students while solving the CPS problems?

Research question 4: What are the relations between the different types of test-taking and problem-solving behaviors and CPS performance?

Methods

Participants

The participants were volunteer undergraduate students ($M_{\text{age}} = 21.50$, $SD_{\text{age}} = 3.03$, $N = 195$) from two Jordanian universities with 15 and 13 faculties, respectively. Students from two faculties took part in the assessment: Arts and Sciences.

Instruments

We adopted a quantitative approach for measuring CPS in the Jordanian higher education context, which is consistent with Molnár et al. (2022). Because the CPS test includes tasks with multimedia components, it requires test-takers to interact with the problem scenarios, thus providing important options for tracking students' test-taking behavior with log data. The basic learning components in KAC and KAP form part of the CPS construct (Molnár & Greiff, 2023). By its very nature, CPS skills represent a crucial educational outcome for the 21st century. It is essential to understand how students learn and subsequently apply knowledge, since it strongly predicts academic success (Schweizer et al., 2013).

CPS was measured with a computer-based test developed within the MicroDYN approach (Greiff & Funke, 2017) and adapted into the Arabic writing style. In MicroDYN, problem environments consist of up to six variables with up to four different types of relations. The problems are embedded in fictitious cover stories, thus eliminating the influence of prior knowledge (e.g., "When you get home in the evening, a young cat is lying on your doorstep. It is exhausted and can barely move. You decide to feed the cat. A neighbor gives you two

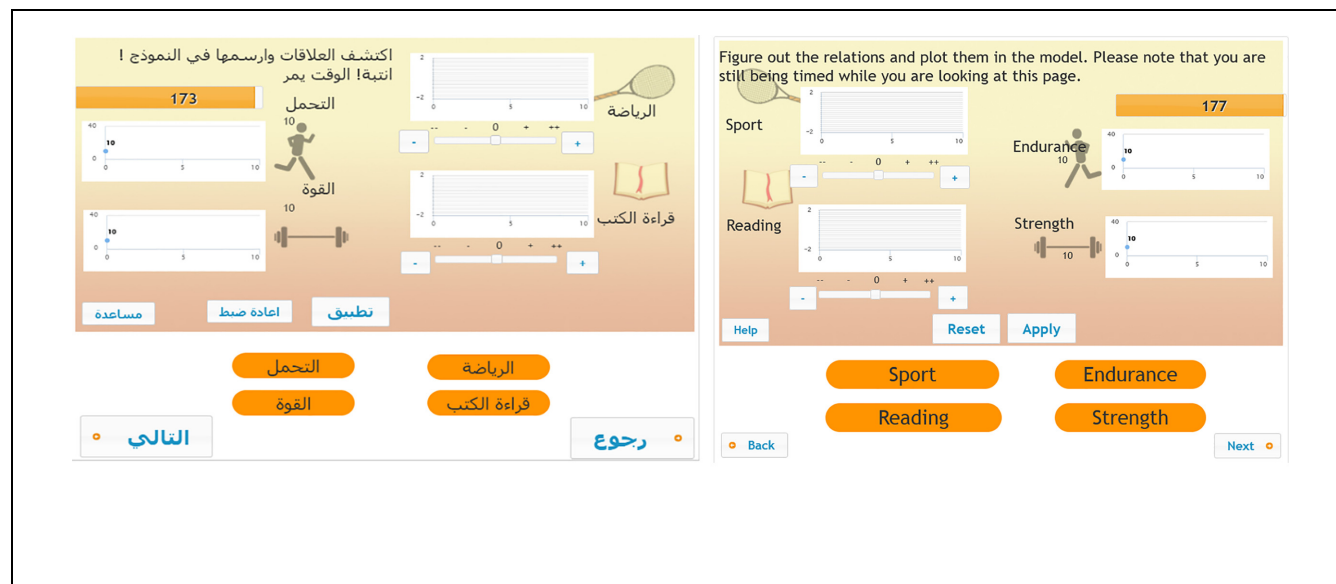


Figure 1. Sample item from the Arabic-language version of the CPS test—KAC phase. In the example, the task is to find out about the effects of sport and reading on endurance and strength. The controllers of the input variables range from “—” (value = -2) to “+ +” (value = $+2$). They are presented on the left side of the problem environment in the English-language version (screenshot on the right) and on the right side in the Arabic-language version (screenshot on the left). The model is shown at the bottom of the figure.

kinds of cat food. Find the relation between the cat food and the cat’s movement/purring”). The test consisted of ten complex MicroDYN problems with different characteristics and different levels of complexity.

Each of the problems consisted of two phases:

First, in the *KAC phase*, participants were expected to explore the structure of the problem scenario by freely operating the system, that is, by manipulating one or more input variables (displayed on the right side according to the Arabic writing style) for no more than 3 min (see Figure 1; sport and reading), and then analyze their effects on the output variables (displayed on the left side according to the Arabic style) (see Figure 1; endurance and strength). Using the “Cat” example noted above, students were asked to ascertain which food affects the cat’s movement and which food influences its purring. Please note that each kind of food in this example only impacted one cat behavior (movement or purring). In parallel, within the 180 s of the KAC phase, they were expected to visualize the relations they detected by drawing lines between the variables on a concept map presented at the bottom of the screen (see Figure 2). The history of the settings is shown on a graph linked to each input and output variable. In practice, each problem scenario has four buttons beyond the adjustment sliders and buttons for the input variables: Help, Apply, Reset, and Next. By clicking on the Reset button, the participant has the

option of deleting all the histories presented on the graphs and setting all the values back to their original values. Each input variable has five stages: $+2$ ($++$), $+1$ ($+$), 0 , -1 ($-$), and -2 ($--$), which can be set using the sliders or buttons ($+$ or $-$) next to the input variables (Figure 1 represents the amount of sport and reading). Their effects on the output variables (the values of endurance and strength) can be tested by clicking on the Apply button. The changes in the output variables are presented in both numerical and graphic formats in the problem scenario. The Next button makes it possible to navigate between the MicroDYN scenarios and its different phases.

Second, in the *KAP phase*, students are expected to use the system in a goal-directed way to reach particular target values (e.g., a given level of movement/purring) of the output variables. To avoid item dependence in this phase, the right concept map is presented at the bottom of the screen. In this part of the problem-solving process, students have no more than 90 s and four trials (clicking four times on the Apply button) to solve the problem, that is, to reach the target values of the output variables. In the case of the “Cat” problem, students were expected to feed the cat properly to reach given target levels of movement and purring. More generally, Figure 3 provides a screenshot of the KAP phase for a problem with four variables (two input and two output variables) with two direct effects.



Figure 2. Example of problem representation: Drawing relations on a concept map provided onscreen. The English-language version is provided on the right.

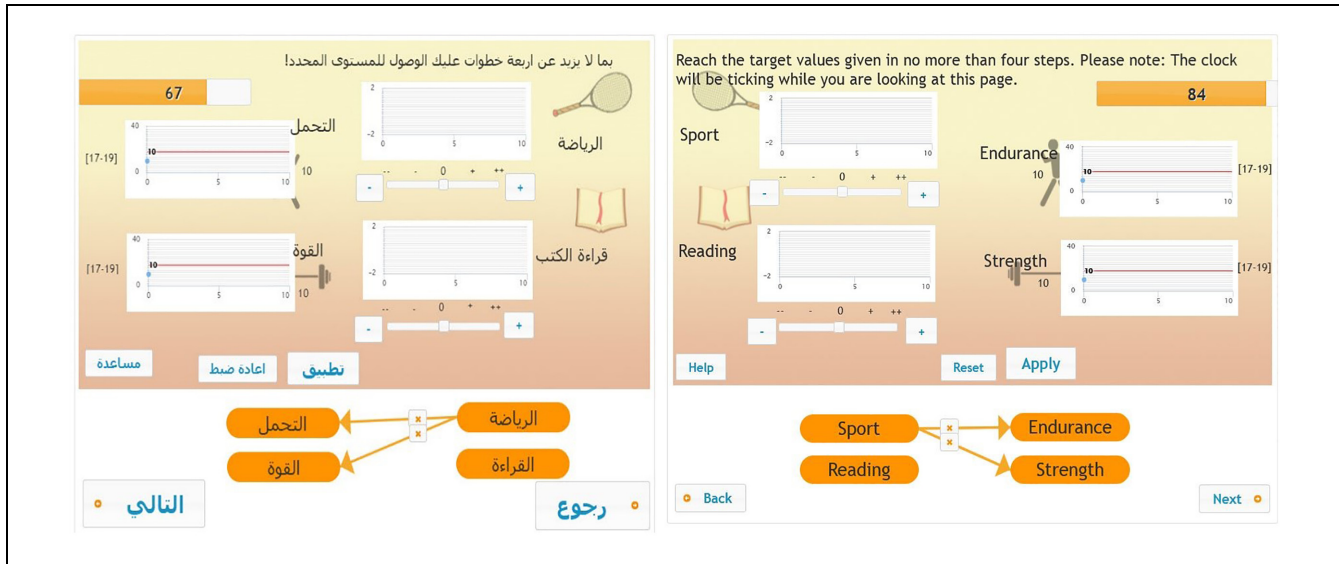


Figure 3. Screenshot of the MicroDYN task “Sport”—KAP phase. The controllers of the input variables range from “— —” (value = −2) to “+ +” (value = + 2). They are presented on the left side of the problem environment in the English-language version (screenshot on the right) and on the right side in the Arabic-language version (screenshot on the left). The right concept map is presented at the bottom of the figure.

Procedure

The items were adapted from the European to the Arabic contexts not only by translating the instructions into Arabic, but also by changing the direction of the items from left to right to make them suitable for the right-to-left reading and writing convention of the Arabic language (see Figures 1 and 3). The complexity of the problem was scaled by the number of variables (input-output; 2-2, 3-2, and 3-3) and the number (2–4) and type (direct or indirect) of relations. According to Beckmann et al. (2017), the rising number of variables and relations increases the difficulty of the CPS problems.

Test Administration. The eDia online assessment platform (<http://www.edia.hu>; Molnár & Csapó, 2019) was used for the test administration. The data collection lasted 45 min at each university’s computer labs. As an achievement indicator, we applied the traditional scoring for both CPS phases (see e.g., Csapó & Molnár, 2017; Fischer et al., 2012; Molnár & Csapó, 2018).

Scoring the Answers and Labeling the Logfiles. If the visualized relations matched the theoretical structure of the problem, students obtained a score of 1. Otherwise, the response was assigned 0 points (for the first phase). In the second phase (KAP), if the problem-solver managed to achieve the target values of the output variables within the given time (90 min) and trial frames (clicking on the Apply button four times), students earned another 1 point or 0 points otherwise. Using the traditional scoring, we generated databases for the analyses for research questions 1 and 2. We had to go beyond traditional scoring to answer research questions 3 and 4. At this point, students’ activity during the problem-solving process was logged and coded based on Molnár and Csapó’s (2018) mathematical model and labeling system, which had been developed based on the effectiveness of the strategy usage and where every trial had been labeled. Students’ problem-solving behavior was defined in each problem situation separately by evaluating all of the trials executed within the same problem. If the problem-solving

behavior followed meaningful regularities, it was labeled a strategy. Three categories were defined within the problem-solving strategies observed: (a) no VOTAT at all, which earned a score of 0 points; (b) partial VOTAT, when VOTAT was only used for some but not all of the input variables, which was assigned a score of 1 point; and (c) full VOTAT, when the VOTAT strategy was used for all the input variables, which garnered a score of 2 points. These scores obtained from the logfiles enabled us to answer research questions 3 and 4.

Data Analysis. The descriptive analyses were executed and bivariate correlations computed by SPSS (for research questions 1, 2, and 4). Confirmatory factor analyses was used to test the underlying measurement model for CPS, assuming two different problem-solving processes, KAC and KAP. These analyses were executed by MPlus (research question 2). We accepted the cut-off values suggested by Hu and Bentler (1999), who indicated that a CFI (Comparative Fit Index) and a TLI (Tucker–Lewis Index) value above .95 and RMSEA (Root Mean Square Error of Approximation) below .06 indicate a good model fit. We used the preferred estimator for categorical variables, Weighted Least Squares Mean and Variance adjusted (WLSMV; Muthén & Muthén, 2010). Latent class analysis (LCA) was used to answer research question 3 and was also executed by MPlus. LCA is a pattern-finding algorithm, which searches for latent classes which share similarly observed variables (Collins & Lanza, 2010). In this study, LCA was used to establish latent classes of students' problem-solving behavior. The quality of the LCA was evaluated with the following fit indices: the Akaike information criterion (AIC), Bayesian information criterion (BIC), and adjusted Bayesian information criterion (aBIC). As regards these fit indices, lower values indicate a better model fit. Entropy was employed to test the accuracy of the classification. The Lo–Mendell–Rubin adjusted likelihood ratio was used to compare the LCA models with different numbers of latent classes (Lo et al., 2001).

Results

Results for Research Question 1 on Testing the Applicability of the CPS Test

When we used the traditional scoring for the CPS problems, the internal consistency of the test was high (Cronbach's $\alpha = .83$). The phase-level reliabilities also proved to be good and acceptable (KAC phase: .83; KAP phase: .65). The test proved to be difficult for the students ($M = 16.8\%$; $SD = 16.7\%$ points), whose achievement was significantly higher in the KAC phase ($M = 25.3\%$; $SD = 25.7\%$ points) than in the KAP phase ($M = 8.1\%$; $SD = 13.0\%$ points; $t = 10.2$,

Table 1. Test and Phase Level Correlations.

Test/ Phase	KAC	KAP
KAC	1.00	
KAP	.453**	1.00
CPS	.925**	.758**

Note. KAC = knowledge acquisition; KAP = knowledge application; CPS = complex problem-solving.
** $p < .01$ level significant.

Table 2. Goodness-of-Fit Indices for Testing the Dimensionality of CPS in Jordan.

Model	χ^2	df	p	CFI	TLI	RMSEA
2-dimensional	234.938	169	.001	0.965	0.961	0.045
1-dimensional	289.945	170	.001	0.936	0.929	0.061

Note. The WLSMV estimator was used in the analyses. df = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation.

$p < .001$). To sum up, using interactive, innovative, third-generation CBA is feasible and reliable in the Jordanian higher education context.

Results for Research Question 2 on the Construct Validity of CPS Measured in the Jordanian Educational Context

The bivariate correlations between the two CPS processes, KAC and KAP, proved to be medium ($r = .45$; see Table 1), indicating measurement of different aspects of CPS.

Confirmatory factor analyses indicated a good fit (see Table 2). A special χ^2 -difference test in Mplus (Muthén & Muthén, 2010) was carried out to compare the one- and two-dimensional models. This test revealed that the two-dimensional model fit the data significantly better (Chi-Square Test for Difference Testing = 55.317, $df = 1$, $p < .001$). Thus, we confirmed the theory and the earlier empirical results based on European and Asian data collections as regards CPS (Wu & Molnár, 2021). CPS is a two-dimensional construct, where the KAC and KAP processes can be distinguished empirically.

Results for Research Question 3 on the Exploration Behavior of the Jordanian University Students While Solving Computer-Based CPS Problems

The efficacy of the exploration strategies was determined and related to the amount of information acquired. If the problem-solver was able to obtain all of the information

Table 3. Percentage of Theoretically Effective and Non-Effective Strategy Use.

Complexity of problem		Percentage (%)	
Number of input and output variables	Number of relations	Effective strategy use	Non-effective strategy use
2-2	2	64.2	35.8
3-2	3	59.8	40.2
3-3	4	56.5	43.5

Table 4. Percentage of No VOTAT, Partial VOTAT, and Full VOTAT Strategy Use.

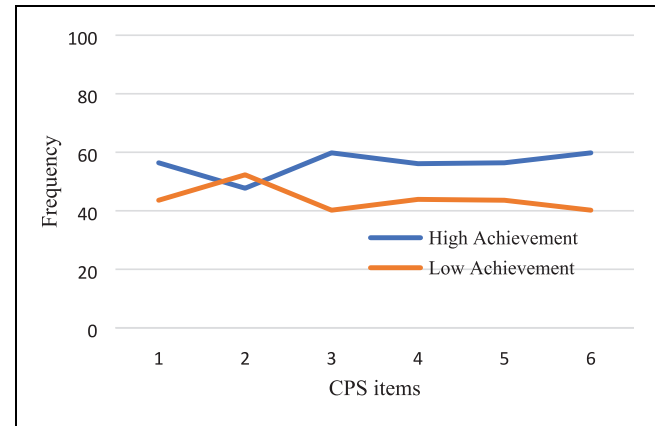
Complexity of problem		No VOTAT (%)	Partial VOTAT (%)	Full VOTAT (%)
Number of input and output variables	Number of relations			
2-2	2	37.4	10.0	52.6
3-2	3	39.2	4.6	56.2
3-3	4	40.0	6.9	53.1

Table 5. Percentage of High and Low Achievers Among the Theoretically Effective Strategy Users During Problem-Solving.

Complexity of problem		Frequency (%)	
Number of input and output variables	Number of relations	High achievement	Low achievement
2-2	2	52.1	47.9
3-2	3	59.8	40.2
3-3	4	57.4	42.6

required to solve the problem correctly in theory, their exploration strategy was defined as theoretically effective (Wu & Molnár, 2022). Contrary to our expectations, based on the results for research question 1, the percentage of theoretically effective strategy use was 56.5% for the more complex problems and 64.2% for the less complex ones (see Table 3).

A large percentage of the Jordanian students employed theoretically effective exploration strategies, including the VOTAT strategy, where the problem-solver manipulates only one input variable systematically while at the same time keeping the other variables unchanged to be able to test the direct effect of the input variables under investigation on the output variables during the problem-solving process. These manipulations allow direct monitoring of changes in output variables to demonstrate the impact of the variable just modified (Molnár & Csapó, 2018). Table 4 summarizes the

**Figure 4.** Problem-solving performance among the theoretically effective strategy users.

percentage of no VOTAT, partial VOTAT, and full VOTAT strategy users. Independently of problem complexity, a majority of the students applied the most effective exploration strategy during the problem-solving process, but, according to the results for research question 1, they were unable to interpret its meaning. That is, at the very end, most of them failed to solve the problems properly.

Results for Research Question 4 on the Relations Between Different Types of Problem-Solving Behavior and Problem-Solving Performance

Only half (52.1%) of the students who applied a theoretically correct strategy made a correct decision as well, solving the easiest problems correctly. This percentage increased to 59.8% with the second sort of complexity before dropping slightly on the most complex problems. Note that the complexity of a problem was defined by the number of variables and the number of relations (Table 5).

Figure 4 displays the percentage of high and low achievers among the theoretically effective strategy users at the task level. The percentage of effective strategy users who correctly completed an item was higher than 50% on most of the items (except item 2). Compared to the relatively low performance for the overall sample (see “Results for research question 1 on testing the applicability of the CPS test” section), the theoretically effective strategy users showed a remarkably better performance.

Problem-solving performance among the theoretically effective strategy users suggests the guessing factor, which indicates a correct solution despite theoretically non-effective strategy usage. This also includes participants who recall a theoretically effective strategy but apply it wrongly and then solve the problem anyway (see Table 6). The guessing factor (indicating those who solve

Table 6. Problem-Solving Effectiveness Among the Theoretically Non-Effective Strategy Users.

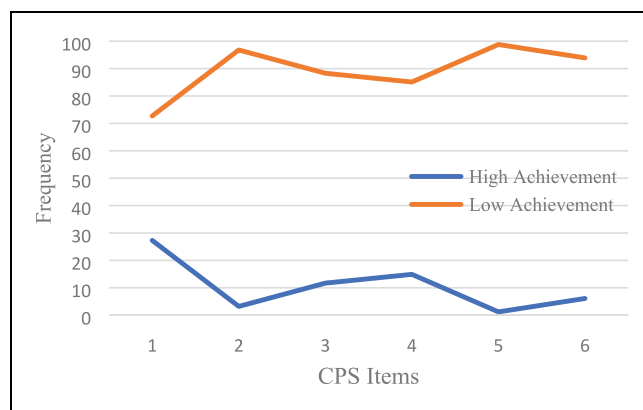
Complexity of problem		Frequency (%)	
Number of input and output variables	Number of relations	High achievement	Low achievement
2-2	2	15.3	84.7
3-2	3	11.7	88.3
3-3	4	7.4	92.6

Table 7. Fit Indices for Latent Class Analyses Monitoring Students' Problem-Solving Behaviour in Uncertain Situations.

Number of latent classes	AIC	BIC	aBIC	Entropy	L-M-R test	p
2	1,504	1,585	1,506	0.967	613	<.001
3	1,446	1,570	1,449	0.931	83	<.05
4	1,440	1,607	1,445	0.944	31	>.05

a problem without an effective strategy) varied from 15.3% to 7.5%, from the least to the most complex tasks. This factor showed the highest effectiveness on the simplest problem (item 1) and then dropped significantly among problems with more than one relation or at least 3 input or 3 output variables. Generally, use of a non-theoretically effective strategy resulted in low achievement on all of the CPS problems (see Figure 5).

After analyzing the performance of theoretically right and theoretically wrong strategy users, we went further to obtain a statistical model of students' problem-solving ability. First, using the tools of latent class analysis and log data analysis to ascertain the use of VOTAT strategies based on students' exploration behavior, we distinguished three qualitatively different VOTAT strategy users. The Akaike, Bayesian and adjusted Bayesian information criterion indices decreased with a growing number of latent classes up to the 4-class solutions. The entropy index reached its maximum value for the 2-class model. However, it was also high for the 3- and 4-class solutions. The Lo-Mendell-Rubin adjusted likelihood ratio test indicated the best model fit for the 3-class model, and it proved to be no longer significant for the 4-class model (see Table 7). Thus, we used the 3-class model—where 93% of the Jordanian students were accurately categorized—to distinguish three qualitatively different class profiles in the further analyses: 50.5% of these students were among the proficient strategy users, who consistently employed VOTAT strategies almost from the very first problem; 18.1% proved to be intermediate explorers, who used VOTAT strategies with lower but still intermediate frequency; and 31.4% were

**Figure 5.** Problem-solving performance among the theoretically non-effective strategy users.

low-level strategy users, who barely made use of VOTAT strategies throughout the assessment process.

Table 8 indicates the problem-solving performance of all three classes of participants (low-level strategy users, intermediate explorers, and expert explorers). The results indicate that all three classes of participants performed better on the easier items (the 2-1 and 2-2 types) than on the more complex problems (the 3-3 type). Furthermore, the results confirmed that VOTAT is the most effective strategy. Problem-solvers that used it had a higher chance of solving a problem correctly, with the exception that the intermediate explorers performed slightly worse than the low-level strategy users on the 3-3 problems.

Discussion

In this study, we used logfile analysis to examine Jordanian undergraduate students' problem-solving behavior. First, we monitored the feasibility and applicability of CBA in the Jordanian educational context. The internal consistency of the CPS tests was high, but the mean achievement was relatively low, indicating that it is difficult for the students to solve interactive problems. Since the internal consistency of the test was high and the phase-level reliabilities also proved to be good and acceptable for both measured phases (KAC and KAP), we can conclude that CBA and innovative online tests are feasible and valid in Jordan at the level and in the context of higher education.

The analyses of the structural stability and validity of the construct being measured confirmed earlier research results obtained in Europe (e.g., Funke, 2001; Wüstenberg et al., 2012) and Asia that CPS is a two-dimensional construct. The processes of the KAC and KAP phases can also be empirically distinguished in the Jordanian context. The bivariate correlation ($r = .45$) between KAC and KAP was consistent with earlier

Table 8. Problem-Solving Performance for Low-Level Strategy Users, Intermediate Explorers and Expert Explorers.

Latent class profiles	Frequency (%)					
	2-2 problems		2-3 problems		3-3 problems	
	High achievement	Low achievement	High achievement	Low achievement	High achievement	Low achievement
Low-level strategy users	32.0	68.0	34.4	65.6	33.9	66.1
Intermediate explorers	33.8	66.2	35.3	64.7	33.3	66.7
Expert explorers	45.4	54.6	46.5	53.5	37.7	62.3

research results, which varied between $r = .14$ and $r = .94$ (Nicolay et al., 2021). The reason for this wide range of correlation coefficients is the use of different problem-solving approaches and CPS assessments to measure KAC and KAP. Since CPS skills are a key competence for educational success, research results on CPS have important implications for filling the gap between students' ability to acquire and then apply that knowledge in uncertain situations, which has become extremely important in the 21st century.

Logfile-based analyses have expanded the scope of previous studies on CPS, especially in the Arabic environment, and enabled us to identify key components of students' problem-solving skills: the way they explore and understand relatively simplified problems and the relationships within the problem. A large number of students showed systematic strategies but failed to solve the problem; that is, the use of a theoretically effective strategy does not always lead to high problem-solving achievement, a finding which confirms research results by De Jong and Van Joolingen (1998), who claim that learners often have trouble understanding data. In contrast, we have detected another relatively large number of students who achieved high performance without collecting all the information necessary to be able to solve the problem correctly; that is, they applied a theoretically non-effective exploration strategy. Beyond guessing, it is more difficult to find a clear explanation for this discrepancy in students' problem-solving behavior. The result is consistent with previous research (e.g., Greiff et al., 2015; Molnár & Csapó, 2018; Vollmeyer et al., 1996) that indicates that high performance is not always in line with the right kind of problem exploration and interpretation. To sum up, the use of a theoretically effective strategy does not always lead to high performance, and, in contrast, high performance does not always indicate the right kind of exploration and interpretation, that is, the application of the right kind of problem-solving strategy.

The analysis explored Jordanian students' problem-solving behaviors in greater depth, focusing on the type of problem exploration and helping us to understand the

reasons behind discrepancies between the high percentage of theoretically right exploration behavior, that is, collecting information, and low problem-solving achievement. One possible explanation is that students did not provide the proper meaning for the information obtained during the first phase of the problem-solving process. Molnár and Csapó (2018) have shown that there is an inverse relation between problem complexity and the probability of strong problem-solving performance without the use of an effective problem-solving strategy. Students' performance was better on problems of medium complexity (2 input and 3 output variables) because they had sufficient experience after solving the first type of problem. On more complex problems (3 input and 3 output variables), students' performance declined despite having sufficient experience in solving problems. As regards the increasing numbers of input variables, output variables and relations between them, the participants experienced greater difficulty (Beckmann et al., 2017). More analyses are required to detect the reasons for the large differences between the expertise level in the exploration and the lower achievement in the decisions made in problem-solving.

Limitations and Conclusions

The study is considered as a small-scale study with 195 participants from two Jordanian universities. Thus, it does not represent the entire university student population in Jordan. A bigger sample size from more universities and faculties is required to obtain a wider view of Jordanian students' problem-solving behavior. Participants had trouble connecting to the Internet during peak use and therefore were occasionally disrupted during the task; all the students used the university system at the same time. This caused some difficulty in retaining access, as some sessions required a high-speed connection. Another limitation stems from the translation and adaptation of the items. Originally, the languages of the items were German and Hungarian. Then, both the Hungarian and German versions were translated into English. After validating the Hungarian,

German and English versions, the test was translated into Arabic by specialist translators for distribution to the Jordanian students. Beyond translating the problem texts and instructions, we changed the direction of the texts on the test to suit the Arabic format, from right to left (earlier versions of the test were produced in left-to-right format). We also changed tables, boxes, images, and all the connecting elements.

The MicroDYN approach was used in the study to assess students' problem-solving abilities with an instrument which is valid and reliable for measurement purposes but uses artificial problems, where the number of variables and relations is limited. Hence, the problem-solving behavior observed in MicroDYN scenarios cannot be generalized to all types of problems we face in everyday life.

Technology has significantly improved the effectiveness of testing procedures by enabling real-time automatic scoring (Dikli 2006), speeding up data processing, facilitating immediate feedback, and revolutionizing the entire assessment process, including the presentation of creative tasks (Csapó et al., 2012). It provides new options for testing and item development. Technology also allows the storing and analysis of contextual data. This new approach is described in the logfile analysis, thus presenting a different type of study (Alrababah & Molnár, 2021).

The study points to the feasibility and construct validity of problem-solving measurements in the Jordanian context. It highlights the importance of explicit development of problem-solving skills and problem-solving strategies as a means of applying knowledge in new contexts in higher education. The findings highlight the importance of developing instructional methods to improve students' CPS skills by enhancing their individual learning strategies.

As regards the educational implications, gaining a better understanding of the differences and similarities in students' problem-solving behavior will not only assist educators in recognizing relevant individual differences more effectively and becoming more sensitive to these differences in learning, but also provide useful input for the design of appropriate training tasks and the training of students to become better problem-solvers.

The results also suggest the need for further investigation to explore the relations between students' cognitive skills and their behavior in problem-solving situations on a larger sample. To sum up, the study has shed light on Jordanian students' problem-solving development from the perspective of their behavior, thus providing a solid basis for further study in the Jordanian context. In addition, it has laid the groundwork for further studies on the measurement of CPS in Jordan and even a cross-national comparison study (see Molnár et al., 2022).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Hungarian National Research, Development and Innovation Fund (grant under the OTKA K135727 funding scheme) and the Hungarian Academy of Sciences (Research Program for Public Education Development of the Hungarian Academy of Sciences, grant KOZOKT2021-16) and by the Humanities and Social Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged. GM is a member of the Digital Learning Technologies Incubation Research Group.

ORCID iDs

Saleh Ahmad Alrababah  <https://orcid.org/0000-0001-7929-2947>

Hao Wu  <https://orcid.org/0000-0002-8560-5046>

Data Availability Statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

- Al-Khayat, M. (2017). Students and instructors' attitudes toward computerized tests in business faculty at the main campus of Al-Blaqa Applied University. *Journal for Research-B (Humanities)*, 31(11), 6.
- Alrababah, S. A., & Molnár, G. (2021). The evolution of technology-based assessment: Past, present, and future. *International Journal of Learning Technology*, 16(2), 134–157.
- Beckmann, J. F., Birney, D. P., & Goode, N. (2017). Beyond psychometrics: The difference between difficult problem solving and complex problem solving. *Frontiers in Psychology*, 8, 1739. <https://doi.org/10.3389/fpsyg.2017.01739>
- Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In: P. A. Frensch & J. Funke: Complex problem solving: The European perspective (pp. 27–63). Hillsdale NJ: Erlbaum.
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.
- Csapó, B., & Funke, J. (2017). Epilogue. In B. Csapó & J. Funke (Eds.), *The nature of problem solving: Using research to inspire 21st century learning* (pp. 265–268). OECD Publishing.
- Csapó, B., Lőrincz, A., & Molnár, G. (2012). Innovative assessment technologies in educational games designed for young

- students. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning* (pp. 235–254). Springer.
- Csapó, B., & Molnár, G. (2017). Potential for assessing dynamic problem-solving at the beginning of higher education studies. *Frontiers in Psychology*, 8, 2022. <https://doi.org/10.3389/fpsyg.2017.02022>
- Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-based assessment of school-readiness and reasoning skills. *Journal of Educational Psychology*, 106(2), 639–650. <https://doi.org/10.1037/a0035756>
- Dede, C. (2010). Comparing frameworks for 21st century skills. In J. Bellance & R. Brandt (Eds.), *21st century skills: Rethinking how students learn* (pp. 51–76). Solution Tree Press.
- De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, 68(2), 179–201. <https://doi.org/10.3102/00346543068002179>
- Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, 5(1), 4–35.
- Dörner, D. (1983). Handeln in Komplexität und Unbestimmtheit. Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität, Bern et al, 19-26.
- Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology*, 8, 1153. <https://doi.org/10.3389/fpsyg.2017.01153>
- Dörner, D., & Wearing, A. J. (1995). Complex problem solving: Toward a (computersimulated) theory. In P. A. Fresch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65–99). Psychology Press.
- Eichmann, B., Goldhammer, F., Greiff, S., Brandhuber, L., & Naumann, J. (2020). Using process data to explain group differences in complex problem solving. *Journal of Educational Psychology*, 112(8), 1546–1562. <https://doi.org/10.1037/edu0000446>
- Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, 4(1), 19–42. <https://doi.org/10.7771/1932-6246.1118>
- Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185–222). Lawrence Erlbaum Associates, Inc.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7(1), 69–89. <https://doi.org/10.1080/13546780042000046>
- Funke, J. (2010). Complex problem solving: A case for complex cognition?. *Cognitive Processing*, 11(2), 133–142. <https://doi.org/10.1007/s10339-009-0345-0>
- Funke, J. (2014). Solving complex problems: Exploration and control of complex systems. In *Complex problem solving* (pp. 185–222). Psychology Press.
- Funke, J. (2021). It requires more than intelligence to solve consequential world problems. *Journal of Intelligence*, 9(3), 38. <https://doi.org/10.3390/jintelligence9030038>
- Funke, J., & Frensch, P. A. (2007). Complex problem solving: The European perspective In D. H. Jonassen (Ed.), *Learning to Solve Complex Scientific Problems* (pp. 25–47).
- Gnaldi, M., Bacci, S., Kunze, T., & Greiff, S. (2020). Students' complex problem solving profiles. *Psychometrika*, 85(2), 469–501. <https://doi.org/10.1007/s11336-020-09709-2>
- Greiff, S., & Funke, J. (2009). Measuring complex problem solving-the MicroDYN approach. In F. Scheuermann (Ed.), *The transition to computer-based assessment-lessons learned from largescale surveys and implications for testing* (pp. 157–163). Office for Official Publications of the European Communities.
- Greiff, S., & Funke, J. (2017). Interactive problem solving: Exploring the potential of minimal complex systems. In B. Csapó & J. Funke (Eds.), *The nature of problem solving* (pp. 93–105). OECD Publishing.
- 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., & Wöstenberg, S. (2014). Assessment with micro-worlds using MicroDYN: measurement invariance and latent mean comparisons. *European Journal of Psychological Assessment*. <https://doi.org/10.1027/1015-5759/a000194>
- Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log-file analyses as a window into students' minds? A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, 91, 92–105.
- 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. <https://doi.org/10.1016/j.compedu.2018.07.013>
- Griffin, P., McGaw, B., & Care, E. (2012). The changing role of education and schools. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 1–16). Springer Science + Business Media B.V.
- Hollnagel, E. (1993). *Human reliability analysis: Context and control*. Academic Press.
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Leutner, D., Wirth, J., Klieme, E., & Funke, J. (2005). Ansätze zur Operationalisierung und deren Erprobung im Feldtest zu PISA 2000. In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern* (pp. 21–36). VS Verlag für Sozialwissenschaften.
- Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767–778.
- 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>
- Mikulincer, M. (1989). Cognitive interference and learned helplessness: The effects of off-task cognitions on performance

- following unsolvable problems. *Journal of Personality and Social Psychology*, 57(1), 129–135. <https://doi.org/10.1037/0022-3514.57.1.129>
- Molnár, G., Arababab, S. A., & Greiff, S. (2022). How we explore, interpret, and solve complex problems: A cross-national study of problem-solving processes. *Heliyon*, 8(1), e08775. <https://doi.org/10.1016/j.heliyon.2022.e08775>
- Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in Psychology*, 9, 302. <https://doi.org/10.3389/fpsyg.2018.00302>
- Molnár, G., & Csapó, B. (2019). *How to make learning visible through technology: The eDia-online diagnostic assessment system* (pp. 122–131). *Proceedings of the 11th International Conference on Computer Supported Education, Heraklion, Crete. Scitepress*.
- Molnár, G., & Greiff, S. (2023). Understanding transitions in complex problem-solving: Why we succeed and where we fail. *Thinking Skills and Creativity*, 50, 101408.
- Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity*, 9, 35–45. <https://doi.org/10.1016/j.tsc.2013.03.002>
- Molnár, G., Greiff, S., Wüstenberg, S., & Fischer, A. (2017). Empirical study of computer based assessment of domain-general dynamic problem solving skills. In B. Csapó & J. Funke (Eds.), *The nature of problem solving: Using research to inspire 21st century learning* (pp. 123–143). OECD Publishing.
- Mousa, M., & Molnár, G. (2019). The feasibility of computer-based testing in Palestine among lower primary school students: Assessing mouse skills and inductive reasoning. *Journal of Studies in Education*, 9(2), 1. <https://doi.org/10.5296/jse.v9i2.14517>
- Mousa, M., & Molnár, G. (2020). Computer-based training in math improves inductive reasoning of 9-to 11-year-old children. *Thinking Skills and Creativity*, 37, 100687. <https://doi.org/10.1016/j.tsc.2020.100687>
- 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. <https://doi.org/10.1037/dev0000764>
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus user's guide*.
- 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, 106594. <https://doi.org/10.1016/j.chb.2020.106594>
- Novick, L. R., & Bassok, M. (2005). *Problem solving*. Cambridge University Press.
- OECD. (2014a). *Results: Creative problem solving – students' skills in tackling real-life problems (Vol. V)*.
- OECD. (2014b). *PISA 2012: Technical report*.
- OECD. (2018). *The future of education and skills. Education 2030*.
- Reeff, J. P., Zabal, A., & Blech, C. (2006). *The assessment of problem-solving competencies: A draft version of a general framework*. German Institute for Adult Education (DIE).
- Schnotz, W., Baadte, C., & Müller, A. (2010). Creative thinking and problem solving with depictive and descriptive representations. In L. Verschaffel, E. De Corte, J. Elen, & T. de Jong (Eds.), *Use of representations in reasoning and problem solving* (pp. 21–45). Routledge.
- Schweizer, F., Wöstenberg, 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. <https://doi.org/10.1016/j.lindif.2012.12.011>
- Van Der Linden, D., Sonnentag, S., Frese, M., & Van Dyck, C. (2001). Exploration strategies, performance, and error consequences when learning a complex computer task. *Behaviour & Information Technology*, 20(3), 189–198. <https://doi.org/10.1080/01449290110047990>
- van Laar, E., van Deursen, A. J., van Dijk, J. A., & de Haan, J. (2020). Determinants of 21st-century skills and 21st-century digital skills for workers: A systematic literature review. *Sage Open*, 10(1), 215824401990017. <https://doi.org/10.1177/215824401990017>
- Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, 20(1), 75–100. https://doi.org/10.1207/s15516709cog2001_3
- Wittmann, W. W., & Hattrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research*, 21(4), 393–409. <https://doi.org/10.1002/sres.653>
- Wu, H., & Molnár, G. (2018). Interactive problem solving: Assessment and relations to combinatorial and inductive reasoning. *Journal of Psychological and Educational Research*, 26(1), 90–105.
- 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*, 36, 1009–1032. <https://doi.org/10.1007/s10212-020-00516-y>
- Wu, H., & Molnár, G. (2022). Analysing complex problem-solving strategies from a cognitive perspective: The role of thinking skills. *Journal of Intelligence*, 10(3), 46. <https://doi.org/10.3390/jintelligence10030046>
- Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving. More than reasoning? *Intelligence*, 40, 1–14. <https://doi.org/10.1016/j.intell.2011.11.003>
- Wüstenberg, S., Greiff, S., Molnár, G., & Funke, J. (2014). Cross-national gender differences in complex problem solving and their determinants. *Learning and Individual Differences*, 29, 18–29. <https://doi.org/10.1016/j.lindif.2013.10.006>