Technology, Knowledge and Learning (2022) 27:69–89
https://doi.org/10.1007/s10758-021-09498-2

ORIGINAL RESEARCH

# It is Not Enough to be Smart: On Explaining the Relation Between Intelligence and Complex Problem Solving

**Matúš Grežo**[1] · **Ivan Sarmány-Schuller**[1]

Accepted: 15 February 2020 / Published online: 29 January 2021


**Abstract**
The main aim of this study was to (a) test the construct validity of complex problem solving (CPS); (b) examine the ability to acquire knowledge as a mediator of the relationship between intelligence and CPS performance; and (c) investigate the personal need for structure as a moderator of the relationship between intelligence and knowledge acquisition. A total of 128 participants completed the self-report Personal Need for Structure scale; the Vienna Matrix Test to assess intelligence; and a new multiple complex systems approach method to assess CPS skills. When analyzing the internal structure of CPS, we found that a two-dimensional model consisting of knowledge acquisition and knowledge application best fitted the data. We also found that the relationship between intelligence and CPS performance was partially mediated by the ability to acquire knowledge. Finally, personal need for structure did not moderate the relationship between intelligence and the ability to acquire knowledge. Our results indicate a need to further investigate other cognitive abilities in interaction with contextual situational factors that could additionally explain variance in CPS performance. Moreover, we also highlight the importance of deeper observation of the knowledge application phase of CPS process.

**Keywords** Complex problem solving · Intelligence · Personal need for structure · Knowledge acquisition · Knowledge application

## 1 Introduction

Since societies are becoming increasingly connected, interdependent and dynamic, the ability to solve complex problems is perceived as one of the key competences required in facing complex challenges of today's world (OECD 2005, 2013, 2014). Individuals often encounter new problem situations characterized by uncertainty, connectivity, intransparency/opaqueness, and complexity. In order to solve these situations, individuals need to

✉ Matúš Grežo
matus.grezo@savba.sk

Ivan Sarmány-Schuller
ivan.sarmany-schuller@savba.sk

[1] Institute of Experimental Psychology, Centre of Social and Psychological Sciences, Slovak Academy of Sciences, Dúbravská cesta 9, 841 04 Bratislava, Slovakia

explore and acquire intransparent/opaque information about the problem situation using non-routine cognitive actions (Scherer, Greiff, and Hautmäki 2015). Since traditional IQ tests have been found to be weak predictors of success in real-life complex and non-academic environments (see Rigas and Brehmer 1999), authors have established a new approach, complex problem solving (CPS), to study complex problems using computer-based simulations of real-life complex situations (e.g., Frensch and Funke 1995; Frensch and Sternberg 1991; Funke 2001).

Since the conceptualization of CPS, many works have investigated the relationship between CPS and intelligence, producing contradictory results ranging from non-significant (e.g., Putz-Osterloh 1981) to very strong correlations (e.g., Funke and Frensch 2007). A more recent study by Wüstenberg, Greiff, and Funke (2012) turned from observing the simple relationship between intelligence and CPS as one-dimensional g-factor, and instead investigated intelligence in relation to specific CPS facets. This led to findings that the relationship between intelligence and CPS performance is mediated by the individual's ability to acquire knowledge about the problem. However, a considerable proportion of variance in how individuals acquire knowledge about the problem remained unexplained, which highlighted the importance of investigating the influence of other cognitive demands affecting the relationship between intelligence and knowledge acquisition. Additionally, assessing CPS as a multi-dimensional factor brought other contradictory results about the construct validity of the CPS process (e.g., Greiff and Fischer 2013; Schweizer, Wüstenberg and Greiff 2013; Wüstenberg, Greiff, and Funke 2012).

The aim of this study was threefold. First, we aimed to test the construct validity of CPS process. Second, we verified Wüstenberg, Greiff, and Funke's (2012) results by investigating the ability to acquire knowledge about the problem as mediator of the relationship between intelligence and knowledge application in the CPS process. Finally, since Wüstenberg, Greiff, and Funke (2012) showed a considerable proportion of unexplained variance in the relationship between intelligence and knowledge acquisition, we decided to extend their mediation model by examining the personal need for structure (PNS) as moderator of the relationship between intelligence and knowledge acquisition.

## 2 Defining Complex Problem Solving

This study uses the definition of complex problem solving proposed by Frensch and Funke (1995). According to them:

> CPS occurs to overcome barriers between a given state and a desired goal state by means of behavioral and/or cognitive, multistep activities. The given state, goal state, and barriers between given state and goal state are complex, change dynamically during problem solving, and are intransparent. The exact properties of the given state, goal state, and barriers are unknown to the solver at the outset. CPS implies the efficient interaction between a solver and the situational requirements of the task, and involves a solver's cognitive, emotional, personal, and social abilities and knowledge (Frensch and Funke 1995, p. 18).

Although it is an older definition, it still reflects the main key aspects of the CPS construct and distinguishes it from other similar psychological concepts, like intelligence. First is that it requires using multistep cognitive or behavioral activities to succeed in complex situations. Funke (2010) emphasizes that CPS requires not only one, but a sequence of cognitive operations such as planning, strategic development, knowledge acquisition and evaluation. Another important aspect of CPS is that it requires an active interaction with the problem. Without active interaction, an individual is not able to acquire intransparent/opaque information and generate knowledge about the problem and thus not capable of solving the problem (Kröner, Plass, and Leutner 2005). Finally, the last characteristic of Frensch and Funke's (1995) definition of CPS is that it involves an individual's cognitive as well as emotional, personal, and social abilities and knowledge. This particular aspect was considered a crucial factor in why general intelligence was unable to explain variation in CPS task performance in early studies (Kluwe, Misiak, and Haider 1991).

## 3 The Process of Complex Problem Solving

The most recent and growing approach in measuring CPS is a multiple complex systems approach (see Greiff, Fischer, Stadler and Wüstenberg 2015; Greiff, Wüstenberg and Funke 2012). Although in the literature this approach is mainly referred to as complex problem solving, it is important to state that it is only a specific area of a broader field of CPS research. In the literature, there are several analogous terms used for referring to this area, including dynamic problem solving (Greiff, Wüstenberg and Funke 2012), interactive problem solving (Fischer et al. 2015), creative problem solving (OECD 2014); or dynamic decision-making (Gonzales, Vanyukov, and Martin 2005).

This approach is based on simple scenarios and everyday situations such as creating new material in a laboratory. Here, an individual actively manipulates input variables (three different compounds: Ladium, Stradium, and Melium), which affect output variables (durability, flexibility, and transparency of material). According to this approach, the CPS process consists of three stages (Greiff, Wüstenberg and Funke 2012). In the first stage, an individual can freely explore the system with no restrictions or goals presented. Individuals can actively manipulate input variables in order to acquire information about the system and create a mental representation of a problem. In the second stage, individuals are asked to draw connections between variables. This phase allows the experimenter to assess whether a proper mental model of the problem was created and the correct information about the relationships between variables was acquired. Finally, in the third stage, individuals have to reach given target values of the output variables by manipulating the input variables. Here, an experimenter assesses the ability to apply acquired knowledge (Funke 2010).

The multiple complex systems approach is specific in several ways. Compared to microworld systems (e.g., Tailorshop: Funke 1983; Multiflux: Kröner, Plass and Leutner 2005), which try to realistically mirror real-world problems with their often extremely great complexity and variable interrelatedness, multiple complex systems involve using a larger number of less complex tasks that do not use multiple independent system structures. In these tasks, the whole problem space has to be explored in order to be successful, whereas some microworld tasks do not allow an active experimentation and exploration of the entire problem space because of their complexity (for further differences between formal framework and the multiple complex systems approach, see Greiff, Wüstenberg and Funke 2012). Herde, Wüstenberg and Greiff (2016) outlined three advantages of using the multiple complex systems approach, namely: greater discrimination between problem solvers with diverse levels of CPS ability; higher degree of reliability of knowledge acquisition and knowledge application measuring; and exclusion of the possibility of irreversible impairment of final CPS scores by single random error at the beginning of the assessment.

## 4 The Dimensionality of CPS Construct: Theory Versus Data

In the literature, the CPS process measured using the multiple complex systems approach, is theoretically and methodologically described as both a two-faceted and three-faceted construct. Considering a three-faceted construct, authors often differ in how they refer to these facets, although they have the same denotation, for instance: rule identification, rule knowledge, rule application[1] (Wüstenberg, Greiff and Funke 2012); exploration, externalization of the mental model, control phase (Funke 2010); exploration strategy, acquired knowledge, control performance (Schweizer, Wüstenberg and Greiff 2013); and information retrieval, model building, forecasting (Greiff, Wüstenberg and Funke 2012). However, since Wüstenberg, Greiff and Funke's (2012) results showed that a two-dimensional model better explained the construct of CPS, in more recent studies, only a two-faceted construct of CPS, consisting of knowledge acquisition and knowledge application, is used (e.g., Baggen et al. 2017; Greiff and Funke 2017; Herborn, Mustafić and Greiff 2017; Kretzschmar, Neubert, Wüstenberg and Greiff 2016; Lotz, Scherer, Greiff and Sparfeldt 2017; Molnár, Greiff, Wüstenberg and Fischer 2017; Scherer, Greiff and Hautmäki 2015).

Despite the consistency of using a two-faceted construct of CPS in recent literature, it is not possible to clearly conclude from the studies investigating the construct validity of CPS whether CPS is a one-, two-, or three-faceted construct. For instance, Schweizer, Wüstenberg and Greiff (2013) tested a three-faceted model encompassing rule identification, rule knowledge, and rule application. Similar to Wüstenberg, Greiff, and Funke (2012), they found a very high correlation between rule identification and rule knowledge, concluding that the two-dimensional model consisting of rule knowledge and rule application fitted the data better compared to the three- and one-dimensional model. Similarly, Greiff et al. (2013a, b) used a MicroDYN method and found that a two-dimensional model showed a better fit for the data than a one-dimensional model. However, when using a MicroFIN method, they did not find any difference between the two- and one-dimensional model fit. Finally, Greiff (2012, in Greiff and Fischer 2013) reported a significantly better data fit for a three-dimensional model compared to a one-dimensional model. This inconsistency in results about the dimensionality of CPS brought us to our first aim, and that is to test the construct validity of CPS process. Since the inconclusive findings provided by extant research do not consistently suggest any hypothesis, we formulate following research question:

*Research question (1): Which of the three CPS models (one-, two-, or three-dimensional) best represents the internal structure of CPS?*

[1] Since the terminology *rule identification*, *rule knowledge*, and *rule application* is very common in the literature (e.g., Funke and Greiff, 2017; Schweizer, Wüstenberg and Greiff, 2013; Wüstenberg, Greiff, and Funke 2012), we decided to use this terminology when addressing a three-faceted model of CPS.

## 5 The Relationship between Intelligence and CPS

Many authors have investigated the relationship between intelligence and CPS skills (see Beckmann and Guthke 1995). Earlier studies by Joslyn and Hunt (1998) and Putz-Osterloh (1981) found non-significant relation between intelligence and CPS. Moreover, Kluwe, Misiak and Haider (1991) reported several other studies, concluding that most of them found only a weak or non-significant relationship. Contrary to this, more recent works found a high correlation between these two constructs (Gonzales, Thomas, and Vanyukov 2005; Greiff and Neubert 2014; Greiff et al. 2013a, b; Kröner, Plass and Leutner 2005; Lotz, Sparfeldt and Greiff 2016; Molnár et al. 2017; Stadler et al. 2017; Wüstenberg, Greiff and Funke 2012; Wüstenberg, Stadler, Hautamäki and Greiff 2014). A number of explanations exist for the inconsistency of the findings on the relation between CPS and intelligence. Stadler et al. (2015) concluded that the zero-correlation found in early studies could be explained by using broader measures of intelligence that include a variety of different cognitive tasks (e.g., factual knowledge). Since more recent studies use more specific latent factors of intelligence that are conceptually closer to the construct of CPS (e.g., fluid intelligence), this relationship was found to be stronger. Another possible explanation of this inconsistency could lie in the conceptualization and measurement of CPS. Kröner, Plass and Leutner (2005) argue that the weak association between intelligence and CPS performance could be the result of non-reliable methods or indicators used for measuring CPS. A more recent study by Lotz, Sparfeldt and Greiff (2016) could support this assumption. They reported a high correlation between a general intelligence test and CPS as measured by MicroDYN, suggesting that it is not a broad conceptualization of intelligence but rather a measurement of CPS that could affect the relationship between intelligence and CPS performance.

Going deeper into the understanding of the relation between intelligence and CPS, Wüstenberg, Greiff and Funke (2012) examined how the intelligence is associated with three CPS facets. They found that knowledge acquisition mediated the relationship between intelligence and CPS performance. An important conclusion of this study was that intelligent individuals showed better CPS skills not because of their reasoning abilities, but rather because they used their intelligence to acquire more knowledge about the problem. The finding that intelligence facilitates knowledge acquisition was supported by two more recent studies. Lotz et al. (2017) found that intelligence facilitated the use of VOTAT (i.e., vary-one-thing-at-a-time) and NOTAT (vary no-thing-at-a-time) problem solving strategies in situations where these strategies were useful. Another study by Wüstenberg et al. (2014) found that fluid intelligence affected CPS performance indirectly through its influence on the use of the VOTAT strategy, while the direct effect of intelligence on CPS performance was weak. These findings suggest that rather than investigating the direct effect of intelligence, research should focus more on how intelligence shapes the process of knowledge acquisition, i.e. which exploration strategies will be used, how many times, or in what order. In fact, examining the mediating effect of knowledge acquisition might explain previous inconsistent findings on the direct relationship between intelligence and CPS performance. Unfortunately, the evidence supporting the indirect effect of intelligence on CPS performance is still limited and further research using various CPS methods is required to support the importance of knowledge acquisition in this relationship. Therefore in this study, we aim to verify these findings by investigating the role of knowledge acquisition as a mediator of the relationship between intelligence and knowledge application. According to Wüstenberg, Greiff and Funke (2012) and Wüstenberg et al. (2014) we hypothesize that:

**Hypothesis (1)**: The ability to acquire knowledge about the problem mediates the relationship between intelligence and problem solving performance.

## 6 The Effect of Personal Need for Structure on the Ability to Acquire Knowledge in Complex Environment

Although both studies by Wüstenberg, Greiff and Funke (2012) and Wüstenberg et al. (2014) showed that intelligence significantly predicted knowledge acquisition, a high amount of variance in knowledge acquisition remained unexplained. Up to now, there have been a limited number of studies trying to further explain this variance. Some studies have examined the effect of working memory (Gonzalez et al. 2005a, b; Schweizer et al. 2013; Bühner, Kröner and Ziegler, 2008; Greiff, Krkovic and Hautamäki 2016), Big-Five personality traits (Greiff and Neubert 2012), or scientific reasoning and learning orientation (Wüstenberg et al. 2014). However, this research often brought contradictory or non-significant findings, suggesting the need for further investigation of other cognitive and personality aspects that could mediate/moderate the relationship between intelligence and ability to acquire knowledge about the problem.

Another stream of research focused on cognitive styles suggests that systematic individual differences in ways of organizing and processing information could affect how one acquires knowledge and creates mental representation of complex situations (Messick 1984). In general, people differ in how they deal with the complexity of the environment. When individuals face new situations, they acquire knowledge through generating hypotheses about the situation and searching for evidence to support these hypotheses (Kruglanski 1989). Although this process is universal, people consistently differ in when they decide to stop the information load and terminate this process. According to Neuberg and Newsom (1993), there are two types of strategies that are used to reduce this information load: avoidance strategies and strategies allowing organizing information into more simplified structure. Importantly, people systematically differ in how and to what extent they use these strategies—a dispositional characteristic known as personal need for structure (PNS). Thompson, Naccarato, and Parker (1989) defined this characteristic as the extent to which persons tend to seek out structured and predictable situations, while also avoiding situations that contain ambiguity and uncertainty.

The research on PNS provides a number of indications that, through avoidance and structuring, PNS affects information processing and knowledge acquisition strategies in complex situations. Firstly, in order to be successful in CPS tasks, an individual has to explore and actively experiment with the environment, which has been shown to be problematic for individuals high in PNS—Sarmány-Schuller (1999) found that these people had problems with active experimentation, they were unwilling to stretch beyond their comfort zone or to change their established ways of thinking, attitudes, and structures. Considering this avoidance strategy within the complex problem solving, individuals high in PNS could be more reluctant to investigate/interact with CPS tasks, resulting in reduced knowledge acquisition about problem and subsequently worse CPS performance, even though their intelligence is high.

Secondly, the research shows that individuals with high PNS use non-systematic and heuristic cognitive processes in order to reduce the complexity and uncertainty of situations (Bar-Tal, Raviv and Spitzer 1999; Grežo and Sarmány-Schuller 2015; Sarmány-Schuller 2000; Schaller, Boyd, Yohannes and O'Brien 1995). Using these simple categories in complex

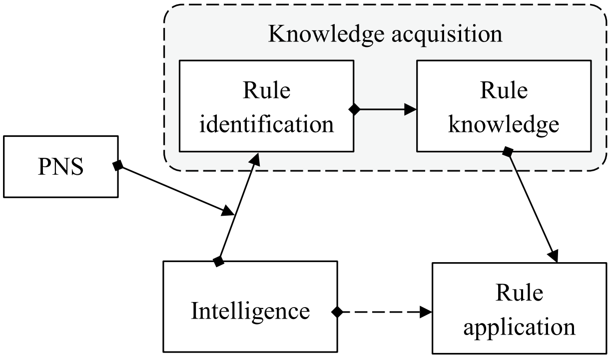


**Fig. 1** Proposed theoretical moderated mediation model of relationships between Intelligence, CPS, and PNS

situations can lead to information loss or incorrect mental representation of situation. In order to reduce complexity, individuals high in need for structure are prone to use inaccurate stereotypes, schemata, or inadequately broad connections (Neuberg and Newsom 1993). This seems to have a great negative effect on performance of tasks that are ambiguous, unstructured and that require creative and divergent thinking (Gocłowska, Baas, Crisp, and De Dreu 2014) which is often the case of complex problem solving tasks. Some other studies support this notion, reporting that PNS negatively affects performance of more creative and divergent intellectual tasks, namely verbal intelligence (Stranovská, Munková, Munk and Sarmány-Schuller 2013), tasks involving monitoring and evaluating several pieces of information at a time (Blais, Thompson and Baranski 2005), or solving word problems with fractions (Svecova and Pavlovicova 2016).

Finally the third important aspect is that high PNS is positively associated with difficulties in integrating multiple pieces of information at the same time, resulting mainly in problems with integrating mathematical information (Sarnataro-Smart 2013), understanding construction of complex spatial structures, and dealing with solving geometrical problem tasks (Wojtowicz and Wojtowicz 2015). This could have particular negative effect in the knowledge acquisition phase where the ability to integrate mathematical information is required and individual is also required to create a mental representation of the relationships between variables. A study by Schultz and Searleman (1998) supports the notion that PNS plays significant role during knowledge acquisition process when the mental representation of the problem is created. They found that in a stress-induced condition, individuals high in PNS tended to solve a given problem using only well-known strategies, i.e. were less flexible when creating a mental representation of a problem. This leaded them to a worse problem solving performance.

Based on the three above mentioned indications, we aim to investigate the personal need for structure as a moderator of the relationship between intelligence and knowledge acquisition. We hypothesize that:

**Hypothesis (2)**: Personal need for structure moderates the relationship between intelligence and knowledge acquisition.

For better understanding of the relationships between all constructs involved in this study, Fig. 1 provides the overall moderated mediation model tested.

# 7 Methods

## 7.1 Research Sample

A total of 128 respondents (109 females, 19 males; age: AM = 20.52; SD = 1.75) participated in this study, all undergraduate students in the social sciences (mainly from social work ≈ 71%, followed by psychology ≈ 29%). The study took place at Constantine the Philosopher University in Nitra, Slovakia, and was conducted during a general psychology course. In this course, participants acquired the basic theoretical knowledge about problem solving as well as other related psychological constructs, like thinking, reasoning, and decision-making. Conducting this study served as an extension of knowledge about a particular field of problem solving. We used opportunity sampling, the participation was voluntary and no refusals were observed during data collecting. Students participated without any financial or course credit reward. The testing procedure was designed so that further tasks were not accessible unless the answers to previous tasks were provided, resulting in no missing responses on the measures.

## 7.2 Procedure

The study and used procedures were designed and conducted according to the standards of the Ethical Codex of Slovak Academy of Sciences and The European Code of Conduct for Research Integrity (ALLEA—All European Academies 2017). The research procedure was programmed in Lazarus software, and testing was entirely computer-based. Before testing, participants were provided basic information about the purpose of the study, the order and nature of tasks as well as the estimated time required for completing these tasks. Moreover, the procedure provided all required information in each particular part of testing in order to fully understand its purpose and nature (e.g. how to explore CPS tasks, how to draw a mental model, how to complete an intelligence test).

The entire process consisted of four parts. In the first part, after reviewing basic information about the purpose of the study, participants provided demographic information (gender, age, study field, achieved education). In the second part, the PNS inventory was administered. In the third part, CPS method was administered. Finally, the fourth part of the procedure assessed intelligence.

After completing the procedure, the programmed software generated a document containing all of the recorded responses. This report consisted of demographics; PNS responses; the number of rounds by which participants explored the system in the first phase of tasks and the numeric values set in each round (which served us to find out whether individuals used a VOTAT analysis); whether the mental model was correctly drawn; the number of rounds the participant needed to reach the targeted goal; and how many (and which) items were correctly solved on an intelligence test. Completing the entire procedure took approximately one hour.

### 7.3 A New CPS Method

To examine CPS ability, we created a method based on a multiple complex systems approach and inspired by a MicroDYN method (Greiff, Wüstenberg and Funke 2012). This method consisted of three tasks differing in difficulty.

In the first task (*The fruit dryer*), the main objective was to manage the temperature and humidity of the fruit dryer. Participants had to determine how three different buttons affected these parameters. The main goal of this task was to set the temperature (T) and humidity (H) to a given target value (T: 40 °C; H: 33%).

In the second task (*Workspace Lighting*), the main objective was to manage three different industrial lamps in a room in order to meet the three parameters (light intensity value: 1000 Lx; discomfort glare: 16 UGLR; color index: 80 Ra) required for optimal workplace conditions.

In the third task (*Treatment of diabetes mellitus*), the main goal was to treat a patient suffering from diabetes mellitus. Participants had to manage the quantity of three different active substances (rosiglitazone, metformin, and acarbose) that influenced three health indicators: blood pressure, blood sugar level, and cholesterol. In this task, participants had to reduce the abnormal values of health indicators (blood pressure: from 11 to 5.5 mmol/l; blood sugar level: from 170 to 125 mmHg; cholesterol: from 7 to 4 mmol/l) and keep these indicators as close as possible to the targeted optimal values over the course of 12 rounds (hypothetical 12 months). This task was the most difficult one, since the health indicators tended to worsen over time (values were constantly rising in each round, known as *eigendynamic*[2]) if the substances were not provided to the patient. Moreover, if systolic blood pressure reached an extremely low value (i.e., less than 70 mmHg), the hypothetical patient died and the task ended prematurely.

The structure of tasks was similar to the MicroDYN method; each one consisted of three main phases. In the first phase, participants had a time-limited opportunity (first task: 60 s; second task: 90 s; and third task: 180 s) to acquire knowledge about the system without an identified objective (rule identification phase). This time restriction was set in a way that it provided a sufficient time to fully explore the system and use VOTAT analysis multiple times for each of the input variables. By adjusting values of input variables (buttons, number of lamps, active substances), one could observe their effect on values of output variables (temperature/humidity, lighting parameters, physical health indicators). After the first phase, participants were asked to draw the mental model of a system using a mouse cursor to create connections between system variables (rule knowledge phase). Participants had the opportunity to reset the drawing whenever needed. After confirming a mental model, the procedure showed a correct model drawing, regardless of whether the participant's model was correctly drawn. In order to provide enough time for rehearsal and recall of the correct model, the drawing was displayed for fifteen seconds to participants. After this, in the third phase, the goal was set and the participant was asked to reach this goal (rule application phase).

When programming our CPS task, we followed the linear equation modeling framework (see Funke 2001). It is based on defining both input and output variables and the type of

[2] An *eigendynamic* is a specific effect that can be programmed into a dependent variable. It refers to a constant increase or decrease of value of this variable itself, independent of other influences or variables. This effect creates an impression that the situation is worsening over time if the proper intervention is not provided.

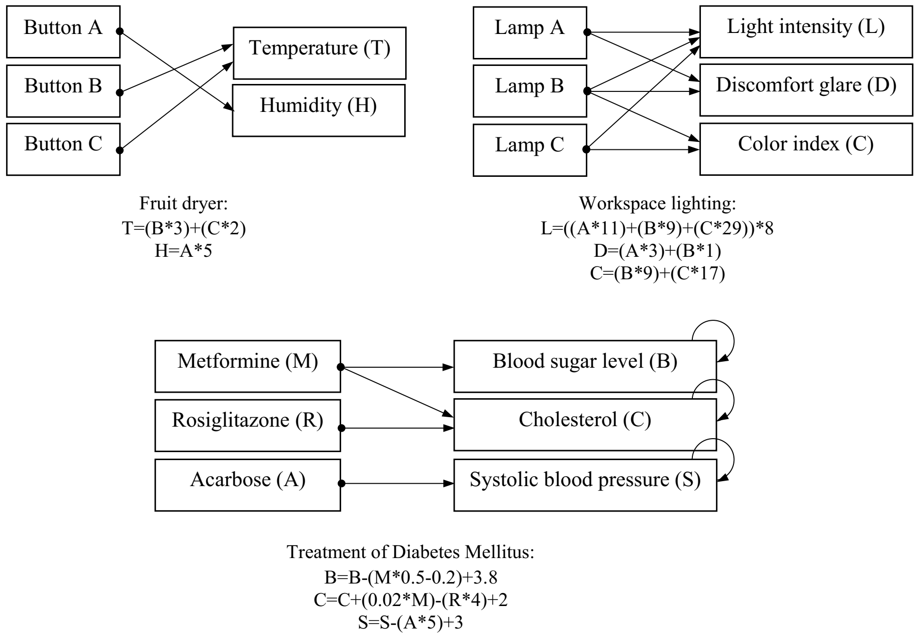


**Fig. 2** Models of relationships between exogenous and endogenous variables used in three CPS tasks. Figure shows how three tasks differed in the number of variables, complexity of relations and also weights with which the exogenous variables affected the endogenous ones. For example, in the Fruit dryer system, one can see that in order to manipulate Humidity (H), an individual had to manipulate button B and C. Moreover, these two buttons had different effect (weight) on humidity—when setting button B to value 1, humidity increased by 3%, compared to button C, which increased humidity only by 2%. Despite less complex relations in the third task compared to the second one, the third task was more difficult because of an eigendynamic used in each dependent variable

relation between them using linear equations. When defining the linear equation, an experimenter has to decide about the number of variables (e.g., the number of buttons in fruit dryer), types of effects between these variables (main effect, side effect, *eigendynamic*), and the strength of these effects. By manipulating these three aspects, an experimenter can adjust the difficulty of the task. When programming three tasks for this study, we manipulated these aspects to vary task difficulty (see Fig. 2 for systems and linear equations used in our CPS tasks).

## 8 Measures

### 8.1 Rule Identification, Rule Knowledge, and Rule Application

As in previous studies (e.g., Wüstenberg, Greiff and Funke 2012; Wüstenberg et al. 2014; Lotz et al. 2017), in order to assess rule identification, we examined whether participants used a VOTAT strategy consistently for each input variable in the task (examining rule identification as a dichotomous variable: 0: not consistently used; 1: consistently used at least once for all input variables in the task). A VOTAT analysis was used when only one input variable was manipulated at a time while other variables were kept constant (e.g., manipulating button A on fruit dryer, while buttons B and C are preserved on an initial level). This strategy is considered to be the best strategy for identifying and understanding causal relations (Tschirgi 1980). The application of the VOTAT strategy is based on the manipulation of only one variable in the experimental design, while others remain unchanged. In such a condition, the effect of the manipulated variable should be fully responsible for resulting changes in dependent variables. Previous studies consistently indicated that using VOTAT strategy leaded to greater knowledge and better problem solving performance compared to other strategies (Vollmeyer, Burns and Holyoak 1996; Kröner et al. 2005; Molnár and Csapó 2018).

When assessing rule knowledge, we examined whether the model drawn was completely correct.

The measure of rule application was the only significant difference between our method and MicroDYN method. Instead of dichotomously exploring whether individual reaches target values in a restricted number of moves (MicroDYN), in our first and second CPS task (*Fruit dryer, Workspace lighting*), we assessed how many rounds participants needed to reach the targeted values. As we mentioned before, in the third task (*Treatment of Diabetes Mellitus*), participants had to reduce abnormal values of health indicators and maintain these indicators as closely as possible to the targeted optimal values throughout 12 rounds. Our first intention was to examine the distance between current and targeted objective values in each round. However, since the third task proved to be difficult to manage (only 23.4% of participants were able to pass all 12 rounds), we decided to examine only how many months (rounds) a participant managed to successfully treat the patient before the task ended prematurely.

### 8.2 Personal Need for Structure

To measure PNS, we used an adapted Slovak version of the Personal Need for Structure scale (Thompson, Naccarato and Parker 1989). It consists of twelve statements assessed on a 6-point Likert scale ranging from strongly disagree to strongly agree. The scale measures two factors: desire for structure and response to lack of structure. A high overall score indicates a strong tendency to prefer certainty and avoid/dislike ambiguous situations. The internal consistency coefficient of PNS in our study showed an acceptable reliability ($\alpha = 0.76$) compared to the good reliability observed in our previous study ($\alpha = 0.84$; Grežo and Sarmány-Schuller 2015). Because of this difference, we calculated McDonald's (1999) omega coefficient, which is more suitable for multidimensional scales. However, the analysis produced a result similar to the Cronbach's alpha ($\omega = 0.75$).

### 8.3 Intelligence

To operationalize intelligence, we used the Vienna Matrix Test (VMT) that was standardized for a Slovak population by Klose, Černochová and Král (2002). This non-verbal test is based on Raven's Progressive Matrices test and consists of 24 items differing in difficulty. Each item consists of 9 visual geometric pictures ($3 \times 3$ matrix) in which the last picture in the bottom right is missing. Participants are expected to identify the pattern between these pictures and correctly fill in the ninth picture with one of the six or eight response pictures provided. The internal consistency coefficient of VMT in our study showed a good reliability ($\alpha = 0.83$).

**Table 1** Descriptive statistics for rule identification and rule knowledge in three CPS tasks

| CPS task | Rule identification | | Rule knowledge | |
|---|---|---|---|---|
| | VOTAT used | VOTAT not used | Correctly drawn | Incor-rectly drawn |
| Fruit dryer | 49.2 | 50.8 | 48.4 | 51.6 |
| Workspace lightning | 59.4 | 40.6 | 53.1 | 46.9 |
| Diabetes mellitus | 68.8 | 31.3 | 38.3 | 61.7 |

**Table 2** Pairwise comparisons of frequencies of rule identification and rule knowledge between three CPS tasks

| Pairwise comparison | Rule identification | | Rule knowledge | |
|---|---|---|---|---|
| | MD | t | MD | t |
| Fruit dryer—Workspace lightning | 0.102 | 2.30* | 0.047 | 1.05* |
| Fruit dryer—Diabetes mellitus | 0.196 | 4.41*** | 0.101 | 2.29 |
| Workspace lightning—Diabetes mellitus | 0.094 | 2.16* | 0.148 | 3.35** |

*MD* Mean Difference, t Student's *t* test value, *$p<0.05$; **$p<0.01$; ***$p<0.001$

# 9 Results

## 9.1 Descriptive Statistics

In order to be able to compare our results with prior studies and conclude about the difficulty of three proposed CPS tasks, we provide descriptive statistics on both CPS and intelligence variables. The mean score in the VMT test for our sample was 98.39 (SD = 12.94). When examining the using of VOTAT strategy, our data showed that only 39% of participants used this strategy consistently in all three CPS tasks. Table 1 shows descriptive statistics for rule identification and rule knowledge in each task separately. When examining how many participants used VOTAT in each task separately, frequencies showed that as tasks proceeded, the number of participants using this strategy gradually increased (see Table 1), while pairwise t-test comparisons showed that this increase was significant (see Table 2). For rule knowledge, the frequency of participants who drew model correctly did not significantly change from first to second task, but this frequency significantly dropped in third task (see Table 1 for frequencies and Table 2 for pairwise comparisons). A possible explanation of this significant drop might be that, compared to the first two tasks, Task 3 contained an autoregressive process which made this task more difficult. Since all three output variables in this task changed their values by themselves, additionally to individuals' manipulations, participants might perceive these independent changes as an outcome of their manipulations, resulting in the creation of incorrect mental representation of the system. Finally for the rule application descriptive statistics, our data showed that the mean number of rounds needed to reach a targeted goal in the first (AM = 20.98; SD = 15.41) and second (AM = 46.46; SD = 41.68) tasks differed significantly ($t = 7.79$; $p < 0.01$). When

**Table 3** Goodness of fit of three complex problem solving models

| Model | $\chi^2$ | df | $p$ | TLI | CFI | RMSEA |
|---|---|---|---|---|---|---|
| 1-DIM | 33.7 | 8 | 0.00 | 0.93 | 0.96 | 0.16 |
| 2-DIM | 5 | 4 | 0.29 | 0.99 | 0.10 | 0.04 |
| 3-DIM | 10.4 | 5 | 0.07 | 0.97 | 0.99 | 0.09 |

1-*DIM* One-dimensional model (CPS as a global factor), 2-*DIM* Two-dimensional model (consisted of knowledge acquisition and knowledge application), 3-*DIM* Three-dimensional model (consisted of rule identification, rule knowledge and rule application)

analyzing rule application in the third task, we found that only 23.4% of participants were able to succeed and manage all 12 rounds without letting the patient die. On average, participants managed to cure the patient for 7.05 months (SD = 4.38).

### 9.2 Confirmatory Factor Analysis

We used confirmatory factor analysis to analyze the construct validity of CPS. We tested one-, two-, and three-dimensional model. When analyzing models using all three tasks, no model showed an acceptable global fit. However, after excluding Task 3 from the analyses, models showed markedly better fit. The best model fit, indicated by the root mean square error of approximation (RMSEA), comparative fit index (CFI) and the Tucker Lewis index (TLI), was shown in a two-dimensional model with rule identification + rule knowledge combined as one factor and rule application as a second factor (see Table 3). All indicators fell within the boundaries recommended by Hu and Bentler (1999). Since the goodness of fit indices indicated that only a two-dimensional model fitted the data well, we used a two-dimensional model consisted of knowledge acquisition (rule identification + rule knowledge) and knowledge application (rule application) in the further analyses. The strength of the association between these facets was found to be moderate and significant ($r = 0.52$; $p < 0.01$).

### 9.3 Intelligence, Personal Need for Structure, and CPS

We investigated the role of knowledge acquisition as a mediator of the relationship between intelligence and knowledge application in the CPS process. When analyzing direct effects, we found that intelligence significantly predicted both knowledge acquisition ($\beta = 0.25$; $p < 0.01$) and knowledge application ($\beta = 0.52$; $p < 0.01$). When we added a path between knowledge acquisition and knowledge application, the effect of intelligence on knowledge application was reduced, although still significant ($\beta = 0.35$; $p < 0.01$), suggesting that intelligence had a significant direct effect on knowledge application, but this effect was also partially mediated by knowledge acquisition. Adding intelligence to the CFA analysis resulted in an acceptable model fit ($\chi^2 = 15.44$; df = 7; $p = 0.03$; TLI = 0.944; CFI = 0.969; RMSEA = 0.08).

Since the considerable amount of variance in knowledge acquisition remained unexplained, we added personal need for structure (PNS) as a moderator of the relationship between intelligence and knowledge application in the model. We firstly ran a regression analysis including both PNS and intelligence as independent variables. Although intelligence was a significant predictor of knowledge acquisition ($\beta = 0.25$; $p < 0.01$), this was not the case for PNS (β = 0.01; $p$ = 0.92). We further investigated the interaction of standardized scores of these two variables on knowledge acquisition and found a non-significant effect (β = 0.14; $p$ = 0.13), suggesting that PNS did not significantly moderate the effect of intelligence on knowledge acquisition.

## 10 Discussion

### 10.1 The Dimensionality of CPS

In two previous studies using MicroDYN, rule identification and rule knowledge facets have been found to be highly correlated, resulting in a better two-dimensional model fit compared to one- and three-dimensional models (Schweizer, Wüstenberg and Greiff 2013; Wüstenberg, Greiff and Funke 2012). In our study, we similarly found that the two-dimensional model fitted the data best and showed a good overall model fit. Other studies using the MicroDYN method also found that the two-dimensional model fitted their data best, suggesting that CPS is a two-dimensional construct (Greiff and Neubert 2014; Greiff et al. 2013a, b). These studies, together with our findings, consistently indicate that CPS is a two-dimensional process consisted of knowledge acquisition and knowledge application. Our findings about the dimensionality of the CPS construct are thus in line with the widely used two-faceted construct in recent research (e.g., Baggen et al. 2017; Greiff and Funke 2017; Herborn, Mustafić and Greiff 2017; Kretzschmar et al. 2016; Lotz et al. 2017; Molnár, Greiff, Wüstenberg and Fischer 2017; Scherer, Greiff and Hautmäki 2015). However, the question remains as to whether such a model can be applied to every complex problem. In our study, we excluded Task 3 from the analyses resulting in significant improvement of model fits. Similar findings appeared in Wüstenberg, Greiff and Funke's (2012) study, where Item 6 showed a low communality on the knowledge application facet. The common specific characteristic of these two tasks is that both of them contain an autoregressive process (*eigendynamic*). Obviously, this characteristic makes CPS tasks much more difficult compared to tasks where an *eigendynamic* is absent. These autoregressive processes complicate the understanding of direct effects between variables and seem to be difficult to control. Like in Wüstenberg, Greiff and Funke's (2012) study, we observed whether VOTAT was used consistently during the knowledge acquisition phase. It seems that this observation is insufficient, especially in *eigendynamic* tasks where the additional examination of how many times a VOTAT was used in knowledge acquisition phase could explain more variability in CPS performance. We emphasize the importance of reassessing and broadening the operationalization of knowledge acquisition indicators in the CPS process.

### 10.2 The Mediating Effect of Knowledge Acquisition on the Relationship Between Intelligence and CPS Performance

A different finding of this study, compared to those of Wüstenberg, Greiff and Funke (2012) and Greiff and Neubert (2014), is that intelligence significantly affected CPS performance both directly and indirectly. This indicated that even when one did not use a VOTAT strategy and did not draw a mental model correctly in the knowledge acquisition phase, she or he was still able to recognize relationships between input and output variables in the last knowledge application phase and solve the problem, although it might take many more rounds or attempts to reach a solution. One possible explanation why our findings differed from those of Wüstenberg, Greiff and Funke (2012) is the difference in how CPS performance was measured during the third application phase. The MicroDYN method explores only whether one is able to reach given target values in a restricted number of moves, resulting in a dichotomous variable (solved or not solved). However, in our method, participants had no such restrictions while solving problem tasks. Instead, we investigated how many rounds or attempts an individual required to reach the target values of the output variables. This method allowed individuals to freely explore the system even in the last application phase. Even when individuals did not understand the system in the first phase, they could explore the system during a number of rounds in the third phase. In other words, our method allowed participants to learn in the application phase without any strict restrictions, allowing a type of digital game-based learning to occur in this phase (see Gros 2015; Prensky 2001; Van Eck 2006). Previous studies have shown that the opportunity to move and act freely in a virtual world promotes active experimentation, strategic thinking, decision making, and other higher cognitive skills such as problem solving (Clark, Tanner-Smith, and Killingsworth 2016; Hainey, Connolly, Stansfield and Boyle 2011; Prensky 2010). This could be beneficial for generating a deeper understanding of complex settings and systems, especially when dealing with multifaceted variables (Gros 2007). Considering our results, it could indicate that the ability to solve complex problems can be promoted using computer-based simulations if time and moves are not restricted. This interpretation could be further supported if we had data about whether an individual who did not use a VOTAT analysis in the first phase tended to use it in the knowledge application phase, i.e. started exploring problem and acquiring knowledge at the last CPS phase. Unfortunately, our method did not report such data. In fact, this limitation also applies to the MicroDYN method; we strongly emphasize the importance of observing the application phase more deeply. This could further explain the differences in findings about the effect of intelligence on the CPS process.

### 10.3 The Moderating Effect of PNS on the Relationship Between Intelligence and Knowledge Acquisition

As in the previous studies (Wüstenberg, Greiff and Funke 2012; Greiff and Neubert 2014; Wüstenberg et al. 2014), although intelligence predicted the ability to acquire knowledge about the problem, a considerable amount of variance in the knowledge acquisition in our study remained unexplained. We assumed that PNS would moderate the relationship between intelligence and knowledge acquisition. However, our results did not support this hypothesis. There are several possible reasons why our assumption was incorrect. Firstly, CPS tasks were new and unique for most of our participants and the main objective in these tasks was to understand relationships between variables, with the only options being to understand them or not to understand them. They cannot be put into a certain scheme in order to reduce its complexity (Neuberg and Newsom 1993). Secondly, in a previous study on PNS, Sarnataro-Smart (2013) found that high-PNS individuals have problems integrating mathematical information. In our CPS tasks, individuals also had to integrate mathematical information, but far less complex. These tasks were not based on finding a correct calculation result. An individual could reach targeted values simply by using a repeated trial-and-error strategy. Even if there would be some differences in how high- and low-PNS individuals carried out calculations, the important is that they did not differ in how they acquired information about problems. A certain analogy could be found in Schultz and Searleman's (1998) study. They found that, when solving a certain problem in a non-stressful condition, both high- and low-PNS individuals were equally sensitive to uncertainty and misperception of complex problem situations and they also tended to apply similar problem solving strategies. However, in stressful conditions, individuals high in PNS tended to develop a mental set that was rigidly used in subsequent problem solving, leading to worse problem solving performance in situations where more optimal strategy was available. This could indicate a third possible explanation of why the effect of PNS in our study was not significant. As Schultz and Searleman (1998) concluded, there exists a contextual activation of personality where motives, heuristics, and schemas are activated through contextual cues. In non-stressful conditions, a personal need for structure could not be activated and therefore did not affect behavior. As we stated before, in our CPS method, participants had enough time to explore systems and to acquire information during the knowledge acquisition phase. Moreover, they were also able to explore the system in the application phase without time pressure or restriction on the number of moves. In such conditions, the effect of PNS could not occur. Unfortunately, we cannot conclude this from our results. Since there is only a limited number of studies investigating the effect of stress on the use of cognitive structuring of information in high- and low-PNS individuals (Bar-Tal, Raviv and Spitzer 1999), we highlight the importance of further investigation on this matter. This could be experimentally tested in future research, where the level of stress (time pressure or other restrictions in the CPS phases) in problem solving should be manipulated. To our best knowledge, such research in the field of CPS has not yet been conducted. There was only one indicative study by Schult et al. (2017) finding that compared to individuals tested in laboratory, those participating at home performed better both in the knowledge acquisition and application phases of CPS, suggesting that laboratory group testing (often associated with higher stress) can have a negative impact on CPS performance.

Investigating the effect of contextual factors, like stress, on CPS performance could have significant implications for practice, especially in professions that are characterized by a great uncertainty, connectivity, and complexity (e.g., firefighter, emergency physician/paramedic, soldier, investor). These professions often deal with complex problem situations that have little structure, are highly stressful (time pressure), and have serious consequences. Previous studies on the effect of PNS on the decision making of healthcare professionals indicated that a higher need for structure could serve as a motivation to start and continue the process of developing a cognitive structure about uncertain situations, leading to a vigilant approach to problem solving and decision-making (Halama and Gurňáková 2014). However, a further investigation is needed to find out whether it also lead to an undesirable rigidity in using these structures which could be associated with biases and errors in reasoning and decision-making (Bar-Tal 2010). Especially in the above mentioned professions, the flexibility in creating and using a certain mental model of a problem situation can be a crucial factor in preventing serious consequences. We highlight the importance of further investigation of contextual situational factors that could affect the ability to acquire knowledge about the problem and choose a proper problem solving strategy.

## 10.4 Study Limitations

Naturally, our study has limitations, mostly in our own method for measuring CPS. We already mentioned some limitations related to the difficulty of the three tasks used. Compared to MicroDYN, where difficulty is slightly increasing from one item to another, the difference in difficulty between our three tasks showed to be much more significant. Especially our third task, characterized by an autoregressive process (*eigendynamic*) implemented for all three output variables, was obviously very difficult to solve. Moreover, input variables in the linear structure equations used for our three tasks had much more weight compared to the MicroDYN method. This could also make our tasks more difficult, since an individual had to operate with higher numbers. When proposing new CPS methods, we suggest creating a larger number of items along with smaller differences in difficulty between these items.

Another issue that could limit the study and affect the results for the effect of PNS is a rather lower reliability of the PNS scale. Although some authors (e.g., Kline 1999) suggest that a cut-off value of 0.7 is suitable when measuring psychological constructs, our findings for the effect of PNS on the relationship between intelligence and CPS should be interpreted with caution. Previous studies have shown that measurement error can result in a reduction in the size of a correlation (see Goodwin and Leech 2006). As Kanyongo, Brook, Blankson and Gocmen (2007) state, researchers often implicitly assume that their measures are perfectly reliable and do not take the reliability effect into account. Unfortunately, this was also the case in our study, and this could result in a lower effect of PNS on the relationship between intelligence and knowledge acquisition. Future studies should emphasize the importance of taking measurement error into account when deciding about required sample size.

## 11 Conclusion

In summary, our results support previous findings about the relation between intelligence and CPS (Greiff and Neubert 2014; Wüstenberg, Greiff and Funke 2012). Our results suggest that CPS is rather a construct consisting of two phases: knowledge acquisition and knowledge application. Although intelligence plays a role in both of these phases, its effect in the first knowledge acquisition phase is more important. Here, intelligence affects whether a correct information acquisition strategy will be used and whether the problem situation will be understood. In addition to intelligence, there seem to be other important cognitive factors that affect which acquisition strategy is used. Our results on the effect of PNS, together with the previous findings of Greiff and Neubert (2014) on the effect of NEO-PI personality traits, suggest that personality factors per se do not play a crucial role in CPS task performance. However, further investigation is needed to examine contextual factors of a CPS situation (e.g., stress condition) that could activate these personality traits and subsequently affect the process of acquiring knowledge and choosing proper problem solving strategies. Finally, we strongly emphasize the importance of more deeply observing the application phase in a CPS process.

**Funding** This work was supported by the grant agency VEGA under Grant no. 2/0116/15 and Grant no. 2/0035/20.

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## References


ALLEA—All European Academies (2017). *The European Code of Conduct for Research Integrity Revised Version*. Berlin: ALLEA—All European Academies.

Baggen, Y., Mainert, J., Kretzschmar, A., Lans, T., Biemans, H. J. A., & Greiff, S. (2017). Complex problems in entrepreneurship education: Examining complex problem-solving in the application of opportunity identification. *Educational Research International, 2017,* 1–13.

Bar-Tal, Y. (2010). When the need for cognitive structure does not cause heuristic thinking: The moderating effect of the perceived ability to achieve cognitive structure. *Psychology, 1,* 96–105.

Bar-Tal, Y., Raviv, R., & Spitzer, A. (1999). The need and ability to achieve cognitive structuring: Individual differences that moderate the effect of stress on information processing. *Journal of Personality and Social Psychology, 77*(1), 33–51.

Beckmann, J. F., & Guthke, J. (1995). Complex problem solving, intelligence and learning ability. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 3–25). Hillsdale: Erlbaum.

Blais, A. R., Thompson, M. M., & Baranski, J. V. (2005). Individual differences in decision processing and confidence judgments in comparative judgment tasks: The role of cognitive styles. *Personality and Individual Differences, 38,* 1701–1713.

Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual–spatial-intelligence and their relationship to problem-solving. *Intelligence, 36*(6), 672–680.

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79–122.

Fischer, A., Greiff, S., Wüstenberg, S., Fleischer, J., Buchwald, F., & Funke, J. (2015). Assessing analytic and interactive aspects of problem solving competency. *Learning and Individual Differences, 39,* 172–179.

Frensch, P. A., & Funke, J. (1995). Definitions, traditions, and a general framework for understanding complex problem solving. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 3–25). NJ: Lawrence Erlbaum Associates.

Frensch, P. A., & Sternberg, R. J. (1991). Skill-related differences in game playing. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 343–381). Hillsdale: Lawrence Erlbaum Associates.

Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? [Issues in problem solving research: Is test intelligence a predictor after all?]. *Diagnostica, 29,* 283–302.

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

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing, 11,* 133–142.

Funke, J., & Frensch, P. A. (2007). Complex problem solving: The European perspective-10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25–47). New York: Lawrence Erlbaum.

Funke, J., & Greiff, S. (2017). Dynamic problem solving: Multiple-item testing based on minimally complex systems. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), *Competence assessment in education: Research, models and instruments* (pp. 427–443). Cham: Springer.

Gocłowska, M. A., Baas, M., Crisp, R. J., & De Dreu, C. K. W. (2014). Whether social schema violations help or hurt creativity depends on need for structure. *Personality and Social Psychology Bulletin, 40*(8), 959–971.

Gonzalez, C., Thomas, R., & Vanyukov, P. (2005a). The relationships between cognitive ability and dynamic decision making. *Intelligence, 33*(2), 169–186.

Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005b). The use of microworlds to study dynamic decision making. *Computers in Human Behavior, 21*(2), 273–286.

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education, 74*(3), 251–266.

Greiff, S., & Fischer, A. (2013). Measuring complex problem solving: An educational application of psychological theories. *Journal for Educational Research Online, 5*(1), 38–58.

Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015). Assessing complex problem-solving skills with multiple complex systems. *Thinking and Reasoning, 21*(3), 356–382.

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013a). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence, 41*(5), 579–596.

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: Using research to inspire 21st century learning* (pp. 93–105). Paris: OECD Publishing.
Greiff, S., Krkovic, K., & Hautamäki, J. (2016). The prediction of problem-solving assessed via microworlds. A study on the relative relevance of fluid reasoning and working memory. *European Journal of Psychological Assessment, 32*(4), 298–306.
Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences, 36,* 37–48.
Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new measurement perspective. *Applied Psychological Measurement, 36*(3), 189–213.
Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013b). Complex problem solving in educational contexts—something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology, 105*(2), 364–379.
Grežo, M., & Sarmány-Schuller, I. (2015). The need for structure and the recognition heuristic in financial decision making. *Československá psychologie: časopis pro psychologickou teorii a praxi, 59*(6), 534–548.
Gros, B. (2007). Digital games in education: The design of games-based learning environments. *Journal of Research on Technology in Education, 40*(1), 23–38.
Gros, B. (2015). Integration of digital games in learning and e-learning environments: Connecting experiences and context. In T. Lowrie & R. JorgensenZevenbergen (Eds.), *Digital games mathematics learning mathematics education in the digital era* (pp. 35–53). Dordrecht: Springer.
Hainey, T., Connolly, T. M., Stansfield, M., & Boyle, E. A. (2011). Evaluation of a game to teach requirements collection and analysis in software engineering at tertiary education level. *Computers and Education, 56*(1), 21–35.
Halama, P., & Gurňáková, J. (2014). Need for structure and big five personality traits as predictors of decision-making styles in health professionals. *Studia Psychologica, 56*(3), 171–179.
Herborn, K., Mustafić, M., & Greiff, S. (2017). Mapping an experiment-based assessment of collaborative behavior onto collaborative problem solving in PISA 2015: A cluster analysis approach for collaborator profiles. *Journal of Educational Measurement, 54*(1), 103–122.
Herde, C. N., Wüstenberg, S., & Greiff, S. (2016). Assessment of complex problem solving: What we know and what we don't know. *Applied Measurement in Education, 29*(4), 265–277.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55.
Joslyn, S., & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. *Journal of Experimental Psychology, 4,* 16–43.
Kanyongo, G. Y., Brook, G. P., Kyei-Blankson, L., & Gocmen, G. (2007). Reliability and statistical power: How measurement fallibility affects power and required sample sizes for several parametric and nonparametric statistics. *Journal of Modern Applied Statistical Methods, 6*(1), 81–90.
Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.
Klose, J., Černochová, D., & Král, P. (2002). *Vídeňský maticový test*. Praha: Testcentrum.
Kluwe, R. H., Misiak, C., & Haider, H. (1991). The control of complex systems and performance in intelligence tests. In H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 227–244). Hillsdale: Lawrence Erlbaum.
Kretzschmar, A., Neubert, J. C., Wüstenberg, S., & Greiff, S. (2016). Construct validity of complex problem solving: A comprehensive view on different facets of intelligence and school grades. *Intelligence, 54,* 55–69.
Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence, 33*(4), 347–368.
Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum.
Lotz, C., Scherer, R., Greiff, S., & Sparfeldt, J. R. (2017). Intelligence in action—Effective strategic behaviors while solving complex problems. *Intelligence, 64,* 98–112.
Lotz, C., Sparfeldt, J. R., & Greiff, S. (2016). Complex problem solving in educational contexts—Still something beyond a "good g"? *Intelligence, 59,* 127–138.
McDonald, R. P. (1999). *Test theory: A unified treatment*. New Jersey: Lawrence Erlbaum Associates.
Messick, S. (1984). The nature of cognitive styles: Problems and promises in educational research. *Educational Psychologist, 19*(2), 59–74.
Molnár, G., Greiff, S., Wüstenberg, S., & Fischer, A. (2017). Empirical study of computer-based assessment of domain-general complex problem-solving skills. In B. Csapó & J. Funke (Eds.), *The nature*

*of problem solving: Using research to inspire 21st century learning* (pp. 125–138). Paris: OECD Publishing.
Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in Psychology, 9,* 302.
Neuberg, S. L., & Newsom, J. T. (1993). Personal need for structure: Individual differences in the desire for simple structure. *Journal of Personality and Social Psychology, 65*(1), 113–131.
OECD. (2005). *The definition and selection of key competencies. Executive summary*. Retrieved from: https://www.oecd.org/pisa/35070367.pdf
OECD. (2013). *PISA 2012 Assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD Publishing.
OECD. (2014). *PISA 2012 Results: Creative problem solving: Students' skills in tackling real-life problems* (Vol. V). Paris: OECD Publishing.
Prensky, M. (2001). *Digital game-based learning*. New York: McGraw-Hill Education.
Prensky, M. (2010). *Teaching digital natives: Partnering for real learning*. California: Corwin—A SAGE Company.
Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg [On the relation between test intelligence and problem solving success]. *Zeitschrift für Psychologie, 189,* 79–100.
Rigas, G., & Brehmer, B. (1999). Mental processes in intelligence tests and dynamic decision-making tasks. In P. Juslin & H. Montgomery (Eds.), *Judgement and decision making: Neo-Brunswikean and process-tracing approaches* (pp. 45–65). Hillsdale: Lawrence Erlbaum Associates Inc.
Sarmány-Schuller, I. (1999). Procrastination, need for cognition and sensation seeking. *Studia Psychologica, 41*(1), 73–85.
Sarmány-Schuller, I. (2000). Need for structure and coping processes. *Ansiedad y Estrés, 6*(1), 39–45.
Sarnataro-Smart, S. (2013). Personal need for structure: Indiscriminate classification systems as barriers to processing mathematical complexity. *Honors Thesis Collection, 110*. http://repository.wellesley.edu/thesiscollection/110
Schaller, M., Boyd, C., Yohannes, J., & O'Brien, M. (1995). The prejudiced personality revisited: Personal need for structure and formation of erroneous group stereotypes. *Journal of Personality and Social Psychology, 68*(3), 544–555.
Scherer, R., Greiff, S., & Hautamäki, J. (2015). Exploring the relation between time on task and ability in complex problem solving. *Intelligence, 48,* 37–50.
Schultz, P. W., & Searleman, A. (1998). Personal need for structure, the einstellung task, and the effects of stress. *Personality and Individual Differences, 24*(3), 305–310.
Schult, J., Stadler, M., Becker, N., Greiff, S., & Sparfeldt, J. R. (2017). Home alone: Complex problem-solving performance benefits from individual online assessment. *Computers in Human Behavior, 68,* 513–519.
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.
Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence, 53,* 92–101.
Stadler, M., Becker, N., Schult, J., Niepel, C., Spinath, F. M., & Greiff, S. (2017). The logic of success: The relation between complex problem-solving skills and university achievement. *Higher Education, 55,* 1–15.
Stranovská, E., Munková, D., Munk, M., & Sarmány-Schuller, I. (2013). Cognitive-individual, linguistic and demographic variables, and syntactic abilities in foreign language. *Studia Psychologica, 55*(4), 273–287.
Svecova, V., & Pavlovicova, G. (2016). Screening the personal need for the structure and solving word problems with fractions. *Springerplus, 5,* 652.
Thompson, M. M., Naccarato, M. E., & Parker, K. E. (1989). *Assessing cognitive need: The development of the personal need for structure and personal fear of invalidity scales*. Halifax: Annual Meeting of the Canadian Psychological Association.
Tschirgi, J. E. (1980). Sensible reasoning: a hypothesis about hypotheses. *Child Development, 51*(1), 1–10.
Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *Educause Review, 41*(2), 1–16.
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.

Wojtowicz, A., & Wojtowicz, B. (2015). The personal need for structure as a factor affecting the understanding and projecting of complex spatial structures. *Technical Transactions Architecture, 11,* 63–72.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving more than reasoning? *Intelligence, 40*(1), 1–14.

Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The role of strategy knowledge for the application of strategies in complex problem-solving tasks. *Technology, Knowledge and Learning, 19*(1–2), 127–146.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.