Journal of Global Education and Research
ISSN 2577-509X (online) ISSN 2577-5081 (print)

Volume 3 | Issue 2 Article 3

December 2019

Influences of reasoning and achievement motivation on complex problem solving in a new microworld operationalization

Stephan Bartholdy
University of Salzburg, stephan.bartholdy@stud.sbg.ac.at

Ulrike Kipman
Institute of Education and Research, ulrike.kipman@phsalzburg.at

Recommended Citation

Bartholdy, S., & Kipman, U. (2019). Influences of reasoning and achievement motivation on complex problem solving in a new microworld operationalization. *Journal of Global Education and Research, 3*(2), 141-157. https://www.doi.org/10.5038/2577-509X.3.2.1035

Corresponding Author

Stephan Bartholdy, Department of Psychology, Faculty of Natural Science, Hellbrunner Straße 34, 5020 Salzburg, Austria

Revisions

Submission date: Aug. 23, 2018 ; 1st Revision: Jan. 25, 2019; 2nd Revision: Jul. 25, 2019; 3rd Revision: Nov. 11, 2019; Acceptance: Nov. 14, 2019

Influences of Reasoning and Achievement Motivation on Complex Problem Solving in a New Microworld Operationalization

Stephan Bartholdy[1] and Ulrike Kipman[2]

Faculty of Natural Science
University of Salzburg, Austria
[1]stephan.bartholdy@stud.sbg.ac.at

Institute of Education and Research
College of Education Salzburg, Austria
[2]ulrike.kipman@phsalzburg.at

Abstract

Complex Problem Solving (CPS) can be defined as those psychological processes that enable a person to achieve goals under complex conditions, which are characterized by their complexity, connectivity, dynamics, lack of transparency, and polytely. Although many hypothesized influences have previously been tested concerning their relevance for the process of solving complex problems (e.g., general intelligence), results were often found to be rather heterogeneous. As this was found to be partially caused by fundamental differences between measurements of CPS, a new operationalization was used in the present study: Following the Microworld approach, CPS was assessed in the simulation game *Cities: Skylines*, as its aptitude as a Microworld could be justified from a theoretical perspective. A parameter for CPS performance was defined to investigate the following hypotheses: Both reasoning and achievement motivation were expected to be positively correlated to CPS performance. Furthermore, a gender difference favoring male participants was expected. Participants in the present study ($N = 27$) first provided demographic information, then subsequently completed a short test of reasoning and an objective personality test of achievement motivation, and finally were given a mission in a complex scenario implemented in *Cities: Skylines*. The results supported all three hypotheses, indicating significant small to moderate positive relationships of both reasoning and achievement motivation with CPS performance, and a significant gender difference favoring male participants in CPS performance. Furthermore, significant gender differences favoring males were also found for reasoning and achievement motivation. Results are discussed and the Microworld operationalization is evaluated.

Keywords: CPS, intelligence, competence, Cities: Skylines, simulation

Introduction

Living in a world of growing possibilities, we are continuously surrounded by challenging tasks and problems. Apparently simple sequences of decisions, like finding a destination in an unknown city, turn out to be highly complex if examined in detail. Such situations have certain characteristics and require specific cognitive abilities similar to those inherent in much more challenging situations, like managing a company, or avoiding the next world war. These characteristics and cognitive processes are subjects of problem-solving research. Concluding from such examples, findings from this research area are highly relevant for a wide range of situations in everyday life. One particularly young topic in this field of research investigates individual differences in how people deal with complex types of problems. What drives psychological research on complex problem solving (CPS), for instance, are questions about the causes for inter- and intraindividual differences in performance when trying to achieve given goals under complex conditions; about the nature of those cognitive and emotional processes that influence our decisions when we find ourselves in new situations; or about strategies we consciously or unconsciously use to solve even the most challenging problems we know in the world.

Literature Review

Before discussing the construct of interest, some key terms from previous research on problem solving ought to be clarified.

Complex Problem Solving

A problem exists when a person does not know how to achieve a current goal. Problem solving therefore corresponds to the cognitive processes that eliminate the barrier between an actual state and a desired goal state (Betsch, Funke, & Plessner, 2011). Concluding from its definition, problem solving is an extremely important set of skills. For this reason, related research gains more and more popularity, and even large-scale assessments like the Programme for International Student Assessment (PISA; see Organisation for Economic Cooperation and Development, 2014) implement both simple- and complex-problem-solving competencies in their assessments. Whereas simple problems are well-defined by a clear set of possible solutions and have a correct solution, complex problems, on the other hand, are ill-defined and have no clear solution.

Characteristics of Complex Problems

More specifically, as formulated by Funke (Dörner & Funke, 2017; Funke, 2001, 2003, 2012), complex problem situations have five characteristic features:

- complexity,
- connectivity,
- dynamics,
- lack of transparency, and
- polytely.

Complexity means that a large quantity of variables is involved in the problem, which therefore need to be considered to be able to solve it. Confronted with many potentially relevant or irrelevant variables, it is necessary to reduce them effectively to find which of them may be important.

Connectivity means that between the large number of variables there also exist many connections. They influence each other in different, not necessarily linear ways. A problem which includes many variables, of which every single one is related to some other variables in different ways, for instance, has a much higher connectivity than a similar problem with the same number of variables, of which each is just influencing one other variable. In the latter problem, the structure of causal relationships could be understood much more easily. In this way, the level of connectivity relates to the (lack of) transparency of problems, as having to understand many interrelations complicates the cognitive process of modelling the problem structure.

Dynamics refer to the attribute of complex problem situations to change over time, both depending on a subject's actions and autonomously. In addition to the large number of interrelated variables, time is a relevant factor for learning processes and for actions to show effects. If dynamic aspects are not considered beforehand, this can result in unexpected consequences of one's own actions. As a special case of dynamics, complex problems can also include eigendynamics, which are autonomous changes not influenced by the subject's interactions with variables. This feature could be induced by variables influencing each other or changing over time in a nonlinear fashion. Eigendynamics, by definition, add an invisible component to the structure of interrelated variables, thereby frustrating the ability to simply draw conclusions from observed changes in the variables. As a consequence, eigendynamics influence both how much knowledge can be gathered about the affected variables, and to what extent these variables can be controlled.

Lack of transparency refers both to the involved variables and goals. All the characteristics described above influence the intransparency of problem situations, because successful problem solving is improbable whenever it is not clear from pure observation if and how relevant variables can be influenced, and which consequences should be expected after taking action. Additionally, goals in complex problem situations are not clearly defined (e.g., find a birthday present for X). As this lack of transparency needs to be eliminated before being able to draw conclusions about possible solutions, the ability to effectively gather information is essential.

Polytely is the simultaneous existence of many goals, which may even be conflicting. For instance, it applies to situations in which the main goal can only be achieved after achieving many smaller goals. Under complex and intransparent conditions, these goals would be conflicting in case of an unexpected negative event. This would interrupt the process of chasing current goals and cause the subject to first care about unexpected problems. Hence, polytely requires evaluating and setting priorities.

A practical example for these characteristics is someone's first interaction with a smart phone: There is an apparently endless amount of menus, symbols, functions, etc. (*complexity*); variables are highly *interconnected* (e.g., changing the system language will affect all menu titles, but not their symbols); the system will change *dynamically* through actions of the user, but also autonomously (e.g., automatically updating weather reports); and at least parts of the underlying causal structure are *not transparent* to the user (e.g., when it shuts down without any apparent reason). Like in this example, research on CPS has shown, when confronted with a complex problem, one needs to interact with it to learn about its underlying causal structure, which is considered necessary for being able to control it.

Measurement Approaches

Approaches to measuring CPS performance can roughly be divided into process-oriented Microworlds and psychometrically oriented approaches, like MicroDYN (Schoppek & Fischer, 2015).

Research on Complex Problem Solving started with Dörner (1980, 1986), who criticized measurements of general intelligence for using overly simple tasks, as compared to the complexity of real-life problems. He proposed the assessment of intelligent behavior in computer-based scenarios which are specifically designed to simulate the characteristics of complex problems in everyday life (Danner, Hagemann, Schankin, Hager, & Funke, 2011). Dörner and his colleagues used so-called Microworlds (see Dörner, Kreuzig, Reither, & Stäudel, 1983) to investigate how subjects acted under complex, dynamic, and intransparent conditions when given a goal which could only be achieved by controlling parts of the system's structure. In their most popular study, participants were given the instruction to manage a small German city called *Lohhausen*, in order to provide the best possible conditions for the city and its population (Hussy, 1998). As this Microworld included more than 2.000 variables, high complexity was assured, thereby also requiring a high-quality operationalization of CPS performance. But as performance in *Lohhausen* was operationalized as a factor assembled from six main criteria (of which some were subjective ratings), its general validity was a critical issue. This definition of a parameter for CPS performance made the interpretation of results very questionable. The reason for using this operationalization could have been that the goal definition–to maximize well-being in the population–was too unclear and open to subjective interpretation (Hussy, 1998). But more recently, some authors (e.g., Danner, Hagemann, Holt et al., 2011; Wittmann & Hattrup, 2004) have shown CPS performance can also be assessed psychometrically valid using Microworlds.

Operationalizations within the MicroDYN approach are computerized tasks which consist of a few linear equations that describe an underlying causal structure between input and output variables. These variables are often represented in numerical values and slide controls. The components can be implemented in different semantic cover stories (e.g., a chemistry laboratory with different substances which can be mixed in different proportions). Within each MicroDYN task, the complex-problem-solving process is split into two phases: (a) the representation phase, in which the participant explores the system (by interacting with it) and is then asked to enter the assumed causal structure into a diagram, followed by (b) the solution phase, in which he or she is shown the correct underlying causal structure and has to control the input variables accordingly to reach given goal values in the output variables. MicroDYN tasks are economic in construction and administration (each task is completed in 5 minutes), the performance measurement is reliable, and tasks can be constructed differing in their semantic cover story and/or causal structure (Schoppek & Fischer, 2015). For instance, Wüstenberg, Greiff, and Funke (2012) found 5% incremental validity of CPS performance in MicroDYN tasks beyond reasoning in predicting school grades. Because of this attribute, they are especially intended for the use as an alternative to common measures of intelligence (Dörner & Funke, 2017).

However, some authors strongly argued against the MicroDYN approach to measure CPS performance (e.g., Funke, 2014; Schoppek & Fischer, 2015). For instance, Funke (2014) criticized the validity of MicroDYN tasks for eliciting other cognitive processes than more complex and naturalistic systems (e.g., Microworlds). It was argued that MicroDYN operationalizations of CPS are lacking (a) *complexity*, as there are only a few variables involved in each task; (b) *connectivity*, as there are also few, mostly linear interrelations between variables; (c) *dynamics*, as there is a restricted number of interactions which participants are given with a system during the short exploration phase, thus not leaving much time for eigendynamics or non-linear (e.g., exponential) influences to show observable effects on the output variables; (d) *intransparency*, as the underlying structure is revealed during the solution phase, when the system has to be controlled towards specific output values; and (e) *polytely*, as MicroDYN tasks may indeed impose more than one goal on participants, but they are neither conflicting nor do they form a sequence of smaller goals required for the achievement of a main goal (Schoppek & Fischer, 2015).

Reasoning

Intelligence is generally considered a fundamentally important set of cognitive skills, as it is empirically shown to be a good predictor of success in many aspects of life (e.g., school grades, see Schmidt & Hunter, 2004). In early research on CPS, intelligence was also expected to be a good predictor of how people deal with complex problems, but the resulting correlations were heterogeneous. In search of reasons for this heterogeneity, and therefore in search of the *real* relationship between measures of intelligence and measures of CPS, many hypotheses were formulated and tested. Extensive research has generally indicated a positive relationship between intelligence and performance in dealing with complex problems (e.g., Beckmann, 1994; Kersting, 1999; Schoppek, 1996; Stadler, Becker, Gödker, Leutner, & Greiff, 2015; Süß, 1996, 1999, 2001), while different interpretations of the relationship would seem plausible. Particularly, it could be assumed that reasoning has an essential role for the processes of gathering information about variables and thereby understanding complex problem situations.

Furthermore, some authors found that domain-specific knowledge influences CPS performance as well (e.g., Elshout, 1987; Raaheim, 1988; Wenke, Frensch, & Funke, 2005). Combining this result with research on intelligence, Elshout and Raaheim both hypothesized there was an inversely U-shaped distribution of the correlation coefficients between CPS performance and intelligence, depending on the extent of domain-specific knowledge. This hypothesized moderation effect is known as the Elshout-Raaheim-Hypothesis (Leutner, 2002). It predicts intelligence would be highly correlated with CPS performance if the extent of domain knowledge was moderate, whereas the correlation would be small for both very little and very much domain knowledge.

Achievement Motivation

Brehmer and Dörner (1993) emphasized that also non-cognitive demands, i.e., coping with ones "more or less strong emotions that are elicited by success or failure" (p. 178), are important factors for the interaction with Microworlds. As further research has shown, some of these non-cognitive influences can be attributed to motivational factors during the problem-solving process. Güss, Burger, and Dörner (2017) illustrated motivational aspects of CPS by interpreting the thinking-aloud protocol of a participant in the *WINFIRE* computer simulation, showing popular theories of motivation were applicable to explain some of the activated needs currently influencing the participant's problem-solving behavior. Vollmeyer and Rheinberg (1999, 2000) found in their studies on motivational aspects of CPS that mastery confidence (i.e., self-efficacy) and incompetence fear were closely related to learning and knowledge acquisition in a complex problem situation. Also, as Kipman (2018) found in children, motivational processes and the problem-solving process in a given situation influence each other. It appears that to overcome barriers in order to reach goals when there is no clear solution, motivation is crucial. A particularly important factor may be achievement motivation, which is defined by PsychInfo as the "need that drives an individual to improve, succeed, or excel" [translated from German] (Schweizer, 2006, p. 224). As concluded by Schweizer, achievement motivation is intrinsic and has three potential sources: the needs to improve one's previous achievements, to chase success (e.g., solving a problem), or to be better than others. Concerning complex problems, it is plausible to assume each of these facets of achievement motivation can actively influence the performance during problem-solving processes.

Gender Differences

In previous research, a relatively robust result among different approaches is a gender difference favoring male participants in complex-problem-solving performance (e.g., Danner, Hagemann, Schankin et al., 2011; Wittmann, Süß, & Oberauer, 1996; Wüstenberg, Greiff, Molnár, & Funke, 2014).

Hypotheses and Aims of the Study

Being one of the most thoroughly analyzed factors in research on CPS, the influence of intelligence (i.e., reasoning) on CPS performance will be examined in the present study. It was hypothesized there is a positive relationship between reasoning and complex-problem-solving performance (H1). Concluding from previous findings, the existence of a moderating effect of domain-specific knowledge on this correlation (Elshout-Raaheim-Hypothesis) will also be investigated. Furthermore, it is expected there is a positive relationship between intrinsic achievement motivation and complex-problem-solving performance (H2). Finally, gender differences favoring male participants in complex-problem-solving performance are expected (H3).

Besides the examination of these hypotheses, another primary aim is to newly operationalize CPS in the computer game *Cities: Skylines*. This city-building simulation meets all necessary criteria which constitute complex problem situations and enable the assessment of CPS in Microworlds.

Methods

In the following section, the sample, material, and procedure of the study are described.

Sample

The present study was conducted at the University of Salzburg, Austria. Participants were recruited online and via registration lists which were handed out in different courses. Students of Psychology received partial course credit for their participation. The sample consisted of 28 participants. One participant was excluded from all analyses due to technical problems while testing. Thus, the final sample consisted of $N = 27$ participants (15 male). The mean age was $M = 23.11$ years ($SD = 3.03$), ranging from 19 to 31 years. They were mostly on the same educational level: $n = 24$ reported to have a High School degree, whereas $n = 1$ reported a degree from Secondary School and $n = 2$ reported a University degree.

Material

To test the hypotheses mentioned in the section above, the following material was used.

Adaptive Matrices Test (AMT)

The AMT (Hornke, Etzel, & Rettig, 2003) is a computer-adapted test of reasoning. Consisting of matrices, it provides for the language-free assessment of general intelligence. The matrices were constructed and selected to be Rasch-homogenous, providing for the adaptive presentation of test items in an economic and reliable way. Reliabilities can be determined beforehand for adaptive versions of the test, and empirically lie between .70 and .86 (Cronbach's α; Hornke et al., 2003). For the present study, the short standard version (S4) of the AMT was used to economically assess reasoning as an estimate of general (fluid) intelligence. This speeded test version consists of 12 linearly presented matrices which need to be solved within 12 minutes. Untransformed raw scores of the Reasoning Parameter Theta (estimate of the person's cognitive ability) were used for analyses, with a reliability of Cronbach's $\alpha = .97$ for the present sample.

DOI: 10.5038/2577-509X.3.2.1035

Objective Achievement Motivation Test (OLMT)

The OLMT (Schmidt-Atzert, 2004) was used to assess different personality aspects of motivation. As a computer-based objective personality test, it is constructed to measure achievement motivation in different miniature situations which challenge the subject's efforts to achieve goals. The OLMT has one version, consisting of three subscales: (a) Intrinsic Achievement Motivation (Baseline), (b) Motivation Through Personal Goals, and (c) Motivation Through Competition. As achievement motivation is the effort to show high performances in situations which can either be succeeded or failed, depending on the quality of performances (Testkuratorium, 2008), this construct (especially intrinsic achievement motivation) was chosen for the examination in relation to complex-problem-solving performance. According to the manual, reliabilities (Cronbach's α) of the OLMT are above .90 for performance measures and between .80 and .90 for the aspiration level in subscale 2 (Schmidt-Atzert, 2004). In the present sample, Cronbach's α was .98 among untransformed raw scores for subscale 1 (intrinsic achievement motivation), which was mainly used for analyses.

Cities: Skylines (CSL)

Cities: Skylines (Colossal Order, 2015a) is a computer-based simulation game about building and managing a city. As declared in the user manual, it "offers endless sandbox play in a city that keeps offering new areas, resources, and technologies to explore, continually presenting the player with new challenges to overcome" (Colossal Order, 2015b, p. 4). This computer game was used as a Microworld, as it meets Funke's (2001, 2012) characteristic features of complex problem situations and Brehmer and Dörner's (1993) criteria for Microworlds, which are part of Funke's specification. These attributes enable the assessment of CPS in participants which are interacting with the Microworld. How these characteristics apply to *Cities: Skylines* will be illustrated in the following examples (see Colossal Order, 2015b; Kooter, 2015).

Complexity: The simulation consists of endless structures (e.g., zones, natural resources, streets, buildings, water and electricity infrastructures), controlling possibilities (e.g., taxes, budgets, loans, traffic control, health, safety and education policies), and parameters (e.g., number of inhabitants, population happiness, pollution, criminality). For instance, when buying a wind turbine, the player can consider its price, weekly budget, noise pollution, production of electricity dependent of its placement, requirement for a connection to the city's electric system, etc.

Connectivity: All changes have consequences in the short- and long-term and the player can estimate some of them from info boxes or conclude from related knowledge in real life. Nevertheless, all such assumptions need to be tested to find causal influences. Connectivity is met as variables in the simulation are highly interrelated, meaning there is not only a vast number of factors in the system (complexity), but each of them is connected to some other factors, influencing each other in different ways. For instance, wind turbines should not be placed near residential zones because their noise pollution upsets people living in the neighborhood, lowering their happiness and the ground value of their property.

Dynamics: The main aspects of autonomous and time-dependent variables are the population's needs and complaints. For instance, zoning demands change autonomously over time (*eigendynamics*), while also partially depending on the player's actions. With the city's population and area extending over time, the water and electricity infrastructure, number of schools, hospitals, cemeteries, etc. which cover the needs of the population in one moment will not be enough at a later moment. In addition, every building or street has a specific lifespan (depending on its use) before it is abandoned and needs to be replaced.

Lack of transparency: Lack of transparency is mainly induced by the complexity and connectivity of the system structure. The vast number of variables and interrelations between them makes the active exploration inevitable. There are also intentionally intransparent features (e.g., irregular death waves independent of the player's actions, or an increasing frequency of fires in the city after the player builds the first fire department [contrary to expectations]).

Polytely: The player is required to constantly check up on demands and complaints of the population because both cannot easily be predicted. With the mission for participants in the present study being to increase the city's population, many influences on the number of inhabitants need to be considered at the same time. Under these conditions, the main goal can only be achieved after many smaller goals (e.g., distributing bus stations strategically for students and workers) are achieved. In case of unexpected problems (e.g., a disease spread from water pollution), the current goals are often conflicting and therefore need to be evaluated quickly to adequately set priorities. Such an unexpected conflict would require the player to stop chasing his/her current goals to analyze and evaluate potential causes and consequences of the new problem, and either apply already known strategies or newly explore practical solutions.

Similar to the popular Microworld *Lohhausen* (Dörner et al., 1983), players in *Cities: Skylines* essentially take the role of the city's mayor, together with all its power and responsibility. For the present study, every participant was given the same scenario with the following preconditions: a small, fully functioning city with a population of 2.600 inhabitants, 50.000 units of money, and the general happiness of the population being 90%. They were given the following mission: *Your mission is to increase the city's population to 5.000 inhabitants. Conditions: The inhabitants should not be unhappy and your bank account should not decrease to 0. Advice: Often, priorities have to be set, so don't forget the mission!* (translation of the German instructions that participants were given in written form by the experimenter).

The mission was accomplished if the population reached 5.000 inhabitants, while having maintained an average happiness level of at least 75%. On the contrary, the mission was failed if (a) the city's population size decreased to 1.000 inhabitants, (b) the bank account reached the value 0, or (c) the maximum playing time of 120 minutes had passed. Derived from the task to increase the city's population, a parameter for complex-problem-solving performance was computed as follows:

$$CPS = \frac{\sum population\ differences\ between\ time\ points}{number\ of\ time\ points - 1} \times \frac{population\ maximum}{population\ goal} \quad (1)$$

The CPS parameter was therefore defined as the averaged population gain over time, weighted by the proportion of the goal (5.000 inhabitants) that was completed. *Cities: Skylines* contains info statistics and diagrams in which players can observe the development of important variables in their city over time. For every participant in the present study, the relevant variable ("population") was exported from CSL using the game modification *CSLMoreGraphs* (User "Ibotaro", 2015).

Participants were instructed on how to navigate through the simulation. Contrary to some other operationalizations, they were also given a short demonstration on a list of basic functions:

DOI: 10.5038/2577-509X.3.2.1035

- placement of streets, buildings, water pumps, and wind turbines;
- zoning (placement of Residential, Commercial and Industrial/Office Zones) and the bulldozer mode;
- structural overview of electricity, water pipes, and garbage disposals; and
- finding info statistics to see what the population needs.

For measurement purposes, they were also instructed not to change the default time settings (as this would produce bias in the measurement points). Without such a profound instruction, gaming-inexperienced participants, by nature, would have performed much worse than gaming-experienced ones, considering the latter group would probably already be familiar with the basic structure of similar computer games. Their domain knowledge would unfairly improve their performance, reducing the influence of the Microworld's characteristics on the problem-solving variable. Hence, the instruction was intended to increase fairness despite different preconditions. In this way, the complex characteristics of the simulation were partially controlled by using a standardized scenario and a specific instruction.

Questionnaire

Participants were asked to provide demographic information (i.e., their age, gender, nationality, and level of education). They also had to rate their amount of prior experience with city-building simulation games on a 4-point scale, ranging from *none* to *very much*. Assessing this information was important to investigate if there were effects of prior knowledge on problem-solving performance in the mission (i.e., gaming-experienced participants being able to explore the Microworld faster than inexperienced participants). This rating served as a definition of domain-specific knowledge in the sense of the Elshout-Raaheim-Hypothesis. On the same questionnaire was a list of 20 in-game symbols with their meanings (e.g., *no electricity*), which participants could consult during the mission. Finally, the difficulty of the mission had to be rated on a 5-point scale ranging from 1 (*very easy*) to 5 (*very difficult*). The experimenter also noted on each questionnaire (a) if the participant completed the mission (success, failure, or participant breakup), and (b) on which time of the day the testing session was taken (morning, afternoon, or evening).

Procedure

The testing procedure took place in a computer laboratory of the Department of Psychology at the University of Salzburg, Austria. Participants were invited individually in groups of up to 3 subjects per testing session. Desks and PCs in this computer lab were arranged in a way which did not allow participants to see each other while completing the tasks. At first, participants were given a short information about the procedure of the study and had to sign an informed consent form. Then they completed the short questionnaire described above. Thereafter, they completed the Adaptive Matrices Test and the Objective Achievement-Motivation Test. Test presentation in the Vienna Test System (Schuhfried, 2011) allowed for the consecutive administration of both computerized tests without any interruption. After a short break, participants received a leaflet explaining the mission in *Cities: Skylines* and a demonstration of the in-game navigation, as well as a verbal instruction on a list of basic functions in the game (see Material Section). Then, each participant was given a laptop to play the mission on. If participants had questions regarding the game, they were only answered if the content was part of the instruction. Approximately every 20 minutes, the experimenter checked up on their city's status. Apart from that, they were left alone to explore the Microworld until they either succeeded, failed, or decided to break up. In any case, they finished their participation in the study by answering the last question on the questionnaire (concerning the difficulty of the mission). This procedure was completed in approximately 90 to 150 minutes, depending on the time spent on the mission.

Results

In this section, the results of the present study were described, starting with the specific hypotheses and ending with further analyses which were calculated post hoc.

Hypotheses

Asked about their prior experience with city-building simulation games, n = 13 participants reported to have *none*, n = 11 reported having had *some experience*, n = 2 reported *much experience*, and n = 1 reported *very much experience*. Regarding mission success in CSL, n = 24 participants succeeded, n = 2 failed to complete the mission, and n = 1 decided to break up early.

Preceding One-Sample Kolmogorov-Smirnov Tests showed violations of the assumption of normality for the Reasoning Parameter Theta (AMT), $D(27) = .224$, $p = .001$. Although other assessed variables did not differ significantly from a normal distribution, a comparison of parametric and corresponding nonparametric analyses revealed substantial differences between their results. Therefore, nonparametric measures were preferred for the following analyses.

H1: There was a significant positive relationship between reasoning and complex-problem-solving performance, $r_\tau = .29$, $p = .044$ (see Table 1). The hypothesized positive relationship was therefore supported by the results. Concerning the influence of domain knowledge, unfortunately, separate correlation coefficients for all levels of prior experience with city-building simulations could not be calculated. Thus, the Elshout-Raaheim-Hypothesis could not be tested in this sample.

Table 1. Nonparametric Correlation of Complex Problem Solving (CPS Performance) With Reasoning (Hypothesis 1), Intrinsic Achievement Motivation (Hypothesis 2), and the Participants' Ratings of the Difficulty of the Mission in Cities: Skylines

Measured Construct	**Correlation with CPS Performance**	
	Correlation Coefficient Kendall's τ	Sig. (2-Tailed)
Reasoning	.29	.044
Intrinsic Achievement Motivation	.27	.047
Reported Difficulty of the Mission	-.59	.000

N = 27

H2: There was a significant positive relationship between intrinsic achievement motivation and complex-problem-solving performance, $r_\tau = .27$, $p = .047$ (see Table 1). This result supports the hypothesized link between CPS and motivation.

H3: A significant difference between male (n = 15) and female (n = 12) participants was found in complex-problem-solving performance, as indicated by a Mann-Whitney Test: The median of CPS performance was significantly higher for males (Mdn = 70.59) than for females (Mdn = 48.52), U = 31.00, p = .003 (see Boxplot Diagram in Figure 1). This result is in line with previous research indicating a gender difference in CPS performance.

DOI: 10.5038/2577-509X.3.2.1035

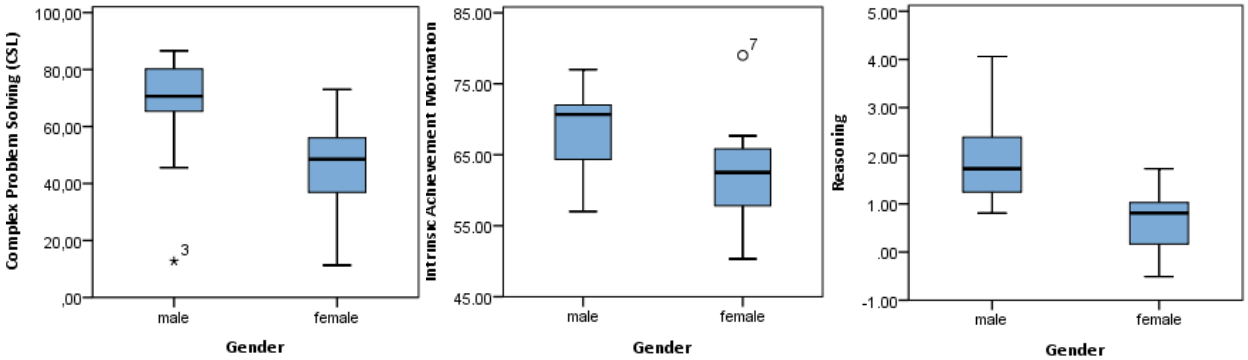

Figure 1. Boxplot diagrams of gender differences in (a) Complex-Problem-Solving Performance in CSL (H3), (b) Intrinsic Achievement Motivation in the OLMT, and (c) Reasoning in the AMT. All three median differences were significant on the $p < .05$ level.

Further Analyses (Post Hoc)

There was a significant negative relationship between CPS performance and the subjective difficulty of the mission, $r_\tau = -.59$, $p < .001$ (see Table 1). This large negative correlation is not surprising, as it indicates the better the participants' CPS performance was, the easier they reported the mission to be.

There was a marginally significant moderate negative relationship between reasoning and the subjective difficulty of the mission, $r_\tau = -.31$, $p = .053$. This indicates the higher the participants' reasoning ability was, the easier they reported the mission to be.

There was a significant positive relationship between motivation through personal goals and motivation through competition, $r_\tau = .46$, $p = .001$. Furthermore, there was a significant positive relationship between motivation through personal goals and aspiration level in the OLMT, $r_\tau = .31$, $p = .023$. These results indicate a high consistency of the personality aspects underlying the facets of motivation addressed in the OLMT. There also was a significant positive relationship between the self-reported amount of prior experience with city-building simulation games and intrinsic achievement motivation, $r_\tau = .32$, $p = .041$.

Mann-Whitney Tests indicated significant differences between male ($n = 15$) and female ($n = 12$) participants in reasoning and intrinsic achievement motivation (see Boxplot Diagram in Figure 1): The median of reasoning was significantly higher for males ($Mdn = 1.73$) than for females ($Mdn = 0.81$), $U = 25.00$, $p = .001$. The median of intrinsic achievement motivation was significantly higher for males ($Mdn = 70.67$) than for females ($Mdn = 60.50$), $U = 40.50$, $p = .014$.

Other analyses of group differences in CPS performance regarding nationality, testing time (morning, afternoon, or evening), and prior experience with city-building simulation games were not significant.

Discussion

The present study examined influences of reasoning, achievement motivation, gender, and the amount of previous experiences with city-building simulations on a parameter for CPS performance that was newly defined and assessed in the Microworld *Cities: Skylines*.

All three hypotheses were supported by the results, because small to moderate positive relationships between both reasoning and CPS performance (H1), and intrinsic achievement motivation and CPS performance (H2) were found in the present sample. Furthermore, gender differences were found in CPS performance (H3) and reasoning, which are also in line with other findings (e.g., see Danner, Hagemann, Schankin et al., 2011; Wittmann et al., 1996).

The result in support of H1 reflects the expected relationship between reasoning as one strong facet of cognitive ability and CPS as another, rather operative and complex one. However, it is unfortunate their relationship could not be calculated and interpreted together with the potential influence of domain-specific knowledge (defined as prior experience with similar computer games). Due to the rather small sample size, it was therefore not possible to test the assumptions of the Elshout-Raaheim-Hypothesis.

It also seems important to mention in this section there were two unexpected results concerning the parameter for intrinsic achievement motivation (OLMT): The moderate positive correlation between the self-reported amount of prior experience with city-building simulation games and intrinsic achievement motivation may be problematic concerning fairness and discriminant validity of the OLMT. It indicates a higher score of intrinsic motivation for more gaming-experienced, as compared to less gaming-experienced participants. Although the reported experience refers to one specific game genre, this result should not be ignored, especially because the author of the test did not find a significant relationship between gaming experience (assessed with a *yes/no* question) and intrinsic achievement motivation (Schmidt-Atzert, 2004). The second issue refers to a significant gender difference favoring male participants in the scores of intrinsic achievement motivation (see Boxplot Diagram in Figure 1), which is also mentioned because the author did not report finding gender differences. Although these results were found in this rather small and homogenous sample, they may still be informative for further research.

Psychometric Evaluation of the Operationalization

Besides fulfilling the criteria for complex problem situations and Microworlds, the newly implemented assessment of CPS performance in *Cities: Skylines* should also suffice the quality criteria for psychological tests (Kubinger, 2003; see Moosbrugger & Kelava, 2012) to be utilised as an assessment instrument. Judging from conditions and results of the present study, the operationalization will be evaluated as comprehensive as possible.

Objectivity is met if the results of a test are independent of the test administrator and the situation, protected from calculation errors, and unambiguous in terms of their interpretation. *Objectivity of Application* can be assumed because of the computerized presentation of the task. However, as the instruction was given verbally (without a verbatim script) during a demonstration of a few basic operations in the game, and as testing sessions varied in the number of participants tested simultaneously (up to 3), the conditions in this procedure were not completely standardized. But this issue could be resolved easily by presenting an instructional video on each participant's computer screen instead of giving a verbal instruction to the group. *Objectivity of Evaluation* (i.e., calculation of scores), on the other hand, was assured by both exporting raw scores directly from the game and calculating the parameter for CPS by computer. Eventually, the O*bjectivity of Interpretation* is not a relevant criterion for the present study, because the resulting parameter was only interpreted relatively for the comparison of participants within the same sample.

Reliability is the accuracy of a test (i.e., the overall consistency of obtained scores). A reliability index (Cronbach's α) of population scores in CSL was calculated for the first 29 measurement time points (out of up to 79 for the participant who lasted the longest in the mission), because this was the last time point that every single participant covered. For this time interval, Cronbach's α was .98 among CSL population scores. The internal consistency was the only reliability estimate that could be calculated in this study, but as CPS is considered a relatively stable construct, it would also be important to assess the test-retest reliability in a longitudinal design in future research.

Validity is met if the results of a measurement reflect only the real construct which was intended to be measured. Above all, high test validity allows one to draw generalized conclusions about behavior under conditions different from those of the testing situation (Moosbrugger & Kelava, 2012). *Content Validity*, in this case, refers to the aptitude of the Microworld's characteristics to elicit complex-problem-solving processes, which could be verified from a theoretical perspective, as well as to the parameter definition for CPS performance in the simulation. This parameter was derived from the task to increase the city's population, making it a pure measure of the outcome of problem-solving processes during the mission. Although the equation is an average measure of performance (i.e., population development) and makes it possible to compare participants regardless of their success in the mission, the interpretation of only one variable, in general, may constitute a validity issue that remains to be clarified in the future. *Face Validity*, however, cannot be evaluated yet because participants of the present study were not informed in detail about the construct and how it was assessed. *Construct Validity* of the current definition of CPS performance can be cautiously estimated from two of the significant results: The CPS parameter was found to be largely independent of reasoning and achievement motivation (discriminant validity). Statements about the concurrent validity of this parameter would be possible by comparing it with results from other operationalizations of CPS (e.g., MicroDYN tasks). *Criterion Validity* is the essential attribute for predicting external behavior from test results and cannot yet be evaluated based on the available information.

Scaling is difficult to evaluate from the collected data: It is met if real differences in the performance of subjects are equally reflected in their test scores. Although the CPS parameter has the great advantage of producing scores on an interval scale, its partial independency from mission success may decrease the scaling quality.

Calibration is not a relevant criterion for this operationalization, as not enough participants have been tested to form a comparative sample. However, assembling a norm sample from tested participants will hopefully be part of further research using this assessment method.

Assuming validity of the operationalization, a moderate *Test Efficiency* could be expected for the task scenario used in the present study, because *Cities: Skylines* can be downloaded and presented very cost-effectively, but the assessment was rather time-consuming. Efficiency could easily be increased by shortening the mission and including more variables in the definition of performance, which can be expected to also increase reliability and validity.

Utility refers to the practical relevance of the assessed construct. Research on the topic appears to gain popularity, which also led to the implementation of CPS tasks in the large-scale assessment PISA (see Organisation for Economic Cooperation and Development, 2014). Concluding from its definition, CPS is an enormously important construct with relevance for countless situations.

Reasonableness of the assessment method can be assumed due to the implementation in a computer game. Being tested in *Cities: Skylines* is expected to not be uncomfortable or problematic for anyone.

Tamper-Resistance is the immunity of test scores from intentional manipulation. As for performance tests in general, it should only be possible for participants to fake a poorer, but not a better performance than they were truly able to achieve. However, as the simulation game is theoretically available for everyone, participants could have already known it (or similar city-building simulations) and be able to perform better than they would have without this experience. To avoid this possibility of manipulation and to increase fairness, detailed instructions were given and participants were asked about their prior experience with similar computer games.

Fairness of the operationalization cannot be evaluated from the available information, because even though there was a significant gender difference, it cannot be concluded yet if the reason was a lack of test fairness or a real gender difference in the construct. Previous research, however, indicates a real gender difference in CPS performance.

Conclusions

Subjectively, the most important part of this study was the theory-based process of evaluating the applicability of *Cities: Skylines* as a Microworld operationalization of CPS. In this evaluation process, many informative conclusions were drawn about how valid the simulation already is in terms of evoking CPS processes, but also about how some of the experienced issues could be resolved (e.g., by standardizing the instruction, or including more variables in the performance parameter). Judging from the theoretical basis of previous research (e.g., see Dörner & Funke, 2017) and from the results of the present study, its design and the operationalization provided for psychometrically satisfying assessment conditions, and furthermore, for demonstrating significant influences of reasoning, achievement motivation, and gender on CPS performance.

Theoretical Implications

Regarding the assessment of complex problem solving, the present study extends the current state of research by adding evidence for the validity of Microworld operationalizations. It could be verified from the definition of the construct that solving the given task in *Cities: Skylines* required CPS skills. The assessment of these skills was a matter of the parameter definition, which was derived from the specified mission. As the parameter variable (i.e., number of inhabitants) was directly recorded over time and exported using a modification of the game (which automatically creates an external datafile), the assessment per se was objective and reliable. Furthermore, the objectivity, reliability, and fairness were increased by using a standardized scenario and a specific instruction, thereby partially controlling the complex features of the simulation. These psychometric properties of the operationalization were evaluated as thoroughly as possible to allow for a judgment about the applicability of *Cities: Skylines* as an assessment method.

Regarding the examined correlates of CPS performance, this study mainly adds to the evidence suggesting links to facets of intelligence and motivation. This suggests both reasoning and achievement motivation play substantial roles in informing and guiding the process of solving problems under complex conditions. In line with previous research, CPS was found to be a matter of cognitive and motivational processes, while it was still largely independent of both. The results therefore support its definition as a distinct construct of cognitive competence.

Practical Implications

As mentioned in the Literature Review, discoveries from this research area are highly relevant for a wide range of situations in everyday life. In light of facing complexity even in the seemingly easiest situations of life, research on complex problem solving has the essential role of finding the underlying processes of our behavior under these conditions. This can be done by using valid operationalizations, such as the Microworld presented in this study. In conclusion, the approach in this study is a plea for validity over efficiency, (i.e., for using complex simulations to assess complex problem solving) rather than using minimally complex scenarios which cannot simulate the complexity of problems in the real world. For this reason, the operationalization of CPS in *Cities: Skylines* was outlined in-depth to facilitate replication studies.

Limitations and Further Research

The main limitation regarding validity and reliability was that the assessment situation was not completely standardized (i.e., regarding the variety of problems which participants faced and the verbal instruction). For instance, it can be argued it was not ensured that all participants would be confronted with the same unexpected events during their time in the mission, because different actions certainly lead to different causal chains of consequences. On the one hand, this feature affects the resulting performance parameter; but on the other hand, it is necessarily inherent in all realistic operationalizations which comply with the established criteria for complex problem situations. The existence of intransparent features and interrelations could possibly penalize the trial-and-error approach which is often taken to solve problems (e.g., when a small error in managing the city could cause extensive damage), meaning imprudence and bad luck could influence the performance. These construct-independent influences, however, were minimized by the relatively long assessment time, in which setbacks could theoretically be compensated for, depending on problem-solving skills and the knowledge previously gathered about the simulation. Hence, it is reasonable that the variety of possible negative consequences should not be minimized for the sake of standardization. In contrast, the verbal instructions on basic functions in the game constitute an issue for objectivity, as they were not completely equal for all participants.

Further empirical research will be needed to replicate the results of this study in larger and more heterogeneous samples, as well as to evaluate the criterion validity of the performance in this operationalization. It would also be informative to examine the test-retest reliability of the operationalization in a longitudinal design, either using the same scenario repeatedly or two or more different scenarios. Another important question regards the influences of personality traits (e.g., risk avoidance, openness to experience, or need for cognition) on CPS performance. Finally, it also remains to be clarified whether there is a moderation effect of domain-specific knowledge on the relationship between reasoning and CPS performance in this Microworld.

References

Beckmann, J. F. (1994). *Lernen und komplexes Problemlösen: Ein Beitrag zur Konstruktvalidierung von Lerntests* [Learning and complex problem solving: A contribution to the construct validation of tests of learning potential]. Bonn, Germany: Holos.

Betsch, T., Funke, J., & Plessner, H. (2011). *Allgemeine Psychologie für Bachelor: Denken - Urteilen, Entscheiden, Problemlösen* [Thinking–Judgement, decision making, problem solving: General psychology for bachelor students]. Wiesbaden, Germany: Springer.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated Microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior, 9*(2-3), 171–184.

Colossal Order. (2015a). Cities: Skylines [Computer software]. Stockholm, Sweden: Paradox Interactive.

Colossal Order. (2015b). *Cities: Skylines: User manual*. Stockholm, Sweden: Paradox Interactive.

Danner, D., Hagemann, D., Holt, D. V., Hager, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). Measuring performance in dynamic decision making. *Journal of Individual Differences, 32*(4), 225–233. doi:10.1027/1614-0001/a000055

Danner, D., Hagemann, D., Schankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence, 39*(5), 323–334. doi:10.1016/j.intell.2011.06.004

Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Games, 11*(1), 87–106.

Dörner, D. (1986). Diagnostik der operativen Intelligenz [Assessment of operative intelligence]. *Diagnostica, 32*(4), 290-308.

Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology, 8,* 1-11. doi:10.3389/fpsyg.2017.01153

Dörner, D., Kreuzig, H. W., Reither, F., & Stäudel, T. (1983). Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität [Lohhausen: On handling uncertainty and complexity]. Bern, Germany: Huber.

Elshout, J. J. (1987). Problem-solving and education. In E. De Corte, H. Lodewijks, & R. Parmentier (Eds.), *Learning & instruction: European research in an international context* (Vol. 1; pp. 259-273). Elmsford, NY: Pergamon Press.

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

Funke, J. (2003). *Problemlösendes Denken* [Problem-solving thinking]. Stuttgart, Germany: Kohlhammer.

Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 682–685). Boston, MA: Springer. doi:10.1007/978-1-4419-1428-6_685

Funke, J. (2014). Analysis of minimal complex systems and complex problem solving require different forms of causal cognition. *Frontiers in Psychology, 5,* 1-3. doi:10.3389/fpsyg.2014.00739

Güss, C. D., Burger, M. L., & Dörner, D. (2017). The role of motivation in complex problem solving. *Frontiers in Psychology, 8*, 1-5. doi:10.3389/fpsyg.2017.00851

Hornke, L. F., Etzel, S., & Rettig, K. (2003). *Manual Adaptive Matrices Test (AMT)*. Mödling, Germany: Schuhfried.

Hussy, W. (1998). *Denken und Problemlösen* [Thinking and problem solving]. Stuttgart, Germany: Kohlhammer.

Kersting, M. (1999). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? Zur Kriteriumsvalidität von Problemlöseszenarien und Intelligenztests* [*Diagnostics and personnel selection with computerized problem-solving scenarios? On the criterion validity of problem-solving scenarios and intelligence tests*]. Göttingen, Germany: Hogrefe.

Kipman, U. (2018). *Problemlösen: Begriff – Strategien – Einflussgrößen – Unterricht - (Häusliche) Förderung* [*Problem solving: Concept – strategies - influencing variables – education - (domestic) promotion*]. Wiesbaden, Germany: Springer.

Kooter, S. de (2015). *Cities Skylines guide: Beginner tips and tricks guide* [Review of the video game *Cities: Skylines*, published by Paradox Interactive, 2015]. Retrieved from https://www.gameplayinside.com/strategy/cities-skylines/cities-skylines-guide-beginner-tips-and-tricks-guide/

Kubinger, K. D. (2003). Gütekriterien [Quality criteria]. In K. D. Kubinger & R. S. Jäger (Eds.), *Schlüsselbegriffe der Psychologischen Diagnostik* [*Key terms of psychological assessment*] (pp. 195–204). Weinheim, Germany: Beltz.

Leutner, D. (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behavior, 18*(6), 685–697. doi:10.1016/S0747-5632(02)00024-9

Moosbrugger, H., & Kelava, A. (2012). Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien) [Quality requirements for a psychological test (Test quality criteria)]. In A. Kelava & H. Moosbrugger (Eds.), *Testtheorie und Fragebogenkonstruktion* [Test theory and construction of questionnaires] (2nd ed.; pp. 7–26). Heidelberg, Germany: Springer.

Organisation for Economic Cooperation and Development. (2014). *PISA 2012 results: Creative problem solving: students' skills in tackling real-life problems (volume V)*. Paris, France: OECD Publishing.

Raaheim, K. (1988). Intelligence and task novelty. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4; pp. 73-97). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmidt-Atzert, L. (2004). *Objektiver Leistungsmotivations-Test OLMT* [Objective Achievement-Motivation Test OLMT]. Mödling, Germany: Schuhfried.
Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of personality and social psychology*, *86*(1), 162. doi:10.1037/0022-3514.86.1.162
Schoppek, W. (1996). *Kompetenz, Kontrollmeinung und komplexe Probleme: Zur Vorhersage individueller Unterschiede bei der Systemsteuerung* [*Competency, locus of control, and complex problems: On the prediction of individual differences in system controlling*]. Bonn, Germany: Holos.
Schoppek, W., & Fischer, A. (2015). Complex problem solving - Single ability or complex phenomenon? *Frontiers in Psychology, 6,* 1-4. doi:10.3389/fpsyg.2015.01669
Schuhfried, G. (2011). Vienna Test System [Computer software]. Mödling, Austria: Schuhfried.
Schweizer, K. (2006). *Leistung und Leistungsdiagnostik* [Performance and performance assessment]. Heidelberg, Germany: Springer.
Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence, 53*, 92–101. doi:10.1016/j.intell.2015.09.005
Süß, H.-M. (1996). *Intelligenz, Wissen und Problemlösen: Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen* [Intelligence, knowledge, and problem solving: Cognitive requirements for the successful interaction with computer-simulated problems]. Göttingen, Germany: Hogrefe.
Süß, H.-M. (1999). Intelligenz und komplexes Problemlösen: Perspektiven für eine Kooperation zwischen differentiell-psychometrischer und kognitionspsychologischer Forschung [Intelligence and complex problem solving: Perspectives for a cooperation between differential-psychometric and cognition-psychological research]. *Psychologische Rundschau, 50*(4), 220-228. doi:10.1026//0033-3042.50.4.220
Süß, H.-M. (2001). Prädiktive Validität der Intelligenz im schulischen und außerschulischen Bereich [Predictive validity of intelligence measures in the educational and non-educational context]. In E. Stern, & J. Guthke (Eds.), *Perspektiven der Intelligenzforschung* [Perspectives of intelligence research] (pp. 109–135). Berlin, Germany: Pabst.
Testkuratorium. (2008). Objektiver Leistungsmotivations-Test (OLMT): TBS-TK Rezension [Objective Achievement-Motivation Test (OLMT): TBS-TK Review]. *Report Psychologie, 33*, 305–306.
User “Ibotaro”. (2015). CSLMoreGraphs [Computer software]. Retrieved from https://steamcommunity.com/sharedfiles/filedetails/?id=577851220
Vollmeyer, R., & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education, 14*(4), 541–554. doi:10.1007/BF03172978
Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction, 10*(4), 293–309. doi:10.1016/S0959-4752(99)00031-6
Wenke, D., Frensch, P. A., & Funke, J. (2005). Complex problem solving and intelligence: Empirical relation and causal direction. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence*: *Identifying the mechanisms of the mind* (pp. 160-187). New York, NY: Cambridge University Press.
Wittmann, W. W., & Hattrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science, 21*(4), 393–409. doi:10.1002/sres.653
Wittmann, W. W., Süß, H.-M., & Oberauer, K. (1996). Determinanten komplexen Problemlösens [Determinants of complex problem solving]. *Berichte des Lehrstuhls Psychologie II der Universität Mannheim, 9*, 1–25. Retrieved from https://www.researchgate.net/profile/Heinz-Martin_Suess/publication/260058458_Determinanten_komplexen_Problemlosens/links/53da25540cf2e38c63365ab4.pdf
Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving–More than reasoning? *Intelligence*, 40(1), 1–14. doi: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. doi:10.1016/j.lindif.2013.10.006