

Lost in Transition – Learning Analytics on the Transfer from Knowledge Acquisition to
Knowledge Application in Complex Problem Solving

Björn Nicolay*¹, Florian Krieger¹, Matthias Stadler², Janice Gobert^{3,4}, & Samuel Greiff¹

¹University of Luxembourg, Faculty of Language and Literature, Humanities, Arts and
Education, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366
Esch-sur-Alzette, Luxembourg, bjorn.nicolay@uni.lu (BN), florian.krieger@uni.lu (FK),
samuel.greiff@uni.lu (SG)

²Ludwig-Maximilians-Universität München, Faculty of Psychology and Pedagogy,
Leopoldstraße 13, 80802 München, Germany, matthias.stadler@uni-muenchen.de

³Rutgers University, Graduate School of Education, 10 Seminary Place, New Brunswick, NJ
08901, US, janice.gobert@gse.rutgers.edu

⁴Apprendis LLC, 57 Village Ct, Berlin, MA 01503, US

Author Note

This research was supported by a grant from the Fonds National de la Recherche Luxembourg
(Luxembourgish National Research Found) [CORE ‘TRIOPS’].

*Corresponding author. ECCS unit, University of Luxembourg, Maison des Sciences Humaines,
11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg. *E-Mail address:*

bjorn.nicolay@uni.lu.

Abstract

Since Complex Problem Solving (CPS) skills represent a key competence for educational success, they are of great relevance for learning analytics. More specifically, CPS serves as a pertinent showcase for addressing a crucial existing gap contemporary education is facing, the gap between students' ability to acquire and subsequently apply knowledge in uncertain situations, which are increasingly important in the 21st century. While the CPS process incorporates both the acquisition and application of knowledge, many earlier studies have focused on identifying the factors relevant for success in knowledge acquisition. Given the dearth of existing research on factors influencing a successful transition between both CPS phases, we investigated the rates of successful and unsuccessful knowledge transition over the course of nine CPS items in a sample of $N = 1,151$ students in 9th grade. Results showed that many participants were unable to transition their knowledge from the acquisition to the application phase, which was presumably due to an inefficient mental model transfer. Furthermore, the likelihood of students being 'lost in transition' was higher in difficult items. Implications are discussed in light of learning analytics, and particularly with regard to the factors to be taken into account by future CPS training programs.

Keywords: complex problem solving, learning analytics, assessment, knowledge acquisition, knowledge application, mental models

Lost in Transition – Learning Analytics on the Transfer from Knowledge Acquisition to
Knowledge Application in Complex Problem Solving

One challenge of education in the 21st century is to ensure that students handle complex and uncertain situations properly. Importantly, this does not only entail the acquisition of knowledge about such a situation, but also its application in a broad range of increasingly complex contexts, in order to solve problems or to make proper decisions (Harris, Krajcik, Pellegrino, & DeBarger, 2019; Kurban, 2018; Laurillard, 2012). However, previous research indicates that the possession of knowledge alone does not guarantee that students are also able to put this knowledge into practice (Charland, Léger, Cronan, & Robert, 2016; Everwijn, Bomers, & Knubben, 1993). Thus, the scientific investigation of the mechanisms behind this transfer of knowledge from acquisition to application has become increasingly relevant in the field of learning analytics (e.g., Saqr, Fors, & Nouri, 2018), which has as one of its primary goals the collection and analysis of students' data to improve their learning process and outcomes (e.g., Zhang, Meng, Ordóñez de Pablos, & Sun, 2019).

In addition, scientific research in educational domains has uncovered a set of skills that are particularly relevant for the students of today. These skills are frequently summarized under the umbrella term '21st century skills' (Adams et al., 2015; Trilling & Fadel, 2009). Mastering these skills can be seen as paramount for the initial educational and future work-related success of students (e.g., Kay & Greenhill, 2011).

Given the established importance for students to improve their performance in various contemporary skills as well as their ability to not only acquire but also to apply knowledge in increasingly complex contexts, the question of how we can use the assessment of students'

performance in a given 21st century skill to gain more insight into the knowledge transition process becomes crucial. In turn, being able to answer this question will provide important implications for future 21st century skill training programs for students, considering both the acquisition as well as the application of knowledge.

In order to study the process of transition between knowledge acquisition and application in light of learning analytics in more detail, this study will focus on complex problem solving (CPS), which has been demonstrated to be a prominent 21st century skill predicting educational success (e.g., Schweizer, Wüstenberg, & Greiff, 2013). On the one hand, due to the ongoing advancement of digitalization, nowadays students are faced with an increasing number of dynamic and complex problem situations that call for more elaborate knowledge acquisition and application skills to be employed under increasingly complex conditions (Charland et al., 2016; Kolfshoten, Lukosch, Verbraeck, Valentin, & de Vreede, 2010; Laurillard, 2012). On the other hand, CPS is generally being assessed in a computer-based way (e.g., Greiff, Wüstenberg, & Avvisati, 2015) and has been included in the most recent cycles of the large-scale Programme for International Student Assessment (PISA; OECD, 2009; 2014), which reflects its relevance for contemporary education and for learning analytics. Importantly, the overarching CPS process stretches over two distinct phases with different individual requirements, knowledge acquisition and knowledge application, thereby reflecting the sequential demands of students in the educational contexts of today to initially acquire knowledge and subsequently put it into practice under increasingly complex conditions (e.g., Novick & Bassock, 2005). Therefore, this study used CPS as a showcase for addressing previously unanswered questions about knowledge transition. Thus, we will be leveraging computer-based student data to advance the knowledge about how a skill that is particularly relevant for education in the 21st century can be fostered.

We will now discuss the construct of CPS and its relevance in the 21st century, and, more importantly, how it represents the educational challenge of knowledge transition, as well as its assessment, while alluding to the potential role of mental models in CPS knowledge transition.

1.1 Complex Problem Solving and Knowledge Transition

One 21st century skill that has received considerable attention in recent educational research is CPS (Greiff et al., 2013; Herde, Wüstenberg, & Greiff, 2016; Schweizer, Wüstenberg, & Greiff, 2013). Complex problem solving can be defined as the ability to solve problems with dynamic, hidden, or intertwined features. Such problems can be encountered on any societal level, from those with global impact such as, for instance, climate change (e.g., Urry, 2008), to individual ones referring to the use of a new smartphone, moving to a new city, or entering high school (e.g., OECD, 2013). In addition, complex problems share several key aspects (Mayer & Wittrock, 2006; Stadler, Niepel, & Greiff, 2019). Firstly, a current state is differing from a desired goal state. In addition, several variables present in the problem space are interrelated ('complexity'), which contain some opaque connections ('intransparency'), and the solver is asked to pursue multiple goals simultaneously ('polytely'). Furthermore, a complex problem encompasses the feature of variables being able to autonomously change over time ('eigendynamics', e.g., Buchner in Frensch & Funke, 1995).

The importance of CPS skills have been uncovered by previous studies in several different domains, for instance with regard to supervisory ratings of job performance of adults (Mainert, Kretzschmar, Neubert, & Greiff, 2015). In addition, previous studies have particularly focused on eliciting the underlying CPS performance indicators in students. Importantly, CPS was discovered to be a significant predictor of educational success (Greiff et al., 2013; Schweizer

et al., 2013), thereby emphasizing its relevance for learning analytics. In addition, CPS inherently captures the increasingly important educational requirement for students to transition their knowledge in a complex context, as its overarching process generally spans over two subsequent phases, *knowledge acquisition* and *knowledge application* (Greiff et al., 2012; Novick & Bassock, 2005).

In the knowledge acquisition phase, a participant is asked to generate knowledge about the existing variable relationships in a given problem space. Thus, the primary goal of this phase is to execute a thorough and fine-grained investigation of the problem space under no or very little time constraints (e.g., Greiff, Fischer, Stadler, & Wüstenberg, 2015). Simultaneously, acquired knowledge of the existing variable relationships in a given problem space is stored as cognitive representations in a mental model (Greiff et al., 2012; Halasz & Moran, 1983, Rickheit & Habel, 1999, Staggers & Norcio, 1993). The process of mental model building starts when a participant has entered the knowledge acquisition phase (i.e., by performing a first variable manipulation in order to detect whether a relationship between two or more variables is present), and is completed when the first phase is terminated (i.e., when the participant believes to have uncovered all existing variable relationships; Funke, 2012).

Subsequently, in the knowledge application phase, solvers are required to put the previously acquired knowledge from their mental model into practice by performing efficient goal-directed variable manipulations in order to reach certain predefined goals (e.g., Funke, 2001). Normally, this has to be accomplished under significant time constraints, such as a very limited number of rounds, before the automated termination of this phase (Greiff et al., 2012).

The two CPS phases can be empirically distinguished according to their different requirements for the solver (e.g., Greiff, Wüstenberg, & Funke, 2012). While during the knowledge acquisition phase, the participant is asked to generate knowledge by almost entirely unrestricted system exploration (i.e., engaging in trial and error), the demands change drastically upon entering the knowledge application phase. Suddenly, the primary goal has shifted to putting the previously acquired knowledge into practice under highly restricting circumstances. Due to this empirical distinction between knowledge acquisition and knowledge application being inherent in CPS, this skill represents a particularly suitable showcase reflecting the challenge of today's education to help students to acquire *and* apply knowledge in increasingly complex situations (Charland et al., 2016; Laurillard, 2012).

Importantly, the relevance of mental models for the knowledge application phase of CPS has already been shown in previous studies (Funke, 2001; Greiff et al., 2012). In addition, the article by Darabi, Nelson, and Seel (2009) describes the improved accuracy of students' mental models in the context of learning complex skills after receiving supportive information, thereby indicating that mental models can be molded and adapted over different points in time; thus, making them a promising candidate for potentially exerting influence on knowledge transition in CPS. That is why we will critically evaluate the potential influence of mental models on a successful or unsuccessful transition between the two CPS phases in our study.

Furthermore, earlier studies are in general agreement about knowledge acquisition and knowledge application being distinct entities, yet significantly correlated (Funke, 2001; Wüstenberg, Greiff, & Funke, 2012). More specifically, the latent correlations between these two phases, as obtained by previous research, range from $r = .14$ to $r = .94$ (Bühner, Kröner, &

Ziegler, 2008; Greiff et al., 2012; 2013; Kröner, Plass, & Leutner, 2005; Neubert, Kretzschmar, Wüstenberg, & Greiff, 2014; Sonnleitner, Keller, Martin, & Brunner, 2013; Wirth, & Klieme, 2003), indicating weak to strong relations between the two phases. However, this broad range of latent correlations is the result of using several CPS assessment tools with different approaches to measure knowledge acquisition and application (for an overview see Herde et al., 2016). Therefore, the selection of an assessment tool that captures the distinct nature of the two phases becomes crucial, in order to be able to draw valid inferences with regard to knowledge transition in CPS.

1.2 Computer-Based Assessment of Complex Problem Solving

Since the advent of computers as a vital tool for scientific research, CPS skills have been assessed mainly in a computer-based way by means of so-called microworlds (Funke, 1993; Funke, 2001; Gobert, Baker, & Wixon, 2015). While, generally, all of these assessment tools capture both phases of CPS, some of them measure knowledge acquisition and knowledge application simultaneously, thereby preventing a clear separation of the respective phases on the measurement level (Funke, 2001). Thus, we will use an assessment approach that clearly distinguishes between the acquisition and application of knowledge (Greiff et al., 2015), which is manifested by a participant receiving an individual performance score for each of the two phases in a given item (Greiff et al., 2012; Sonnleitner et al., 2012).

One particular advantage of assessing students' CPS skills via such computer-based microworlds is that several performance indicators are simultaneously being collected and stored in log files (see also Gobert et al., 2015; Gobert, Sao Pedro, Raziuddin, & Baker, 2013). This process generally happens without the participant being made explicitly aware of it (e.g., Adams

et al., 2015), and is sometimes referred to as stealth assessment (Shute, Wang, Greiff, Zhao, & Moore, 2016). These log files contain more information than mere success or failure of a student in a given microworld, such as, for instance, exactly where and when the participant made a mistake, and how close they ultimately were to achieving a particular goal in the knowledge application phase (e.g., Greiff et al., 2015). Generally, such thorough investigations of performance-related data provide a deep insight into how students approach complex problems, and are frequently performed in the context of learning analytics (Kim, Yoon, Jo, & Branch, 2018; Lim et al., 2019; Nistor & Hernández-García, 2018).

Previous research has already leveraged the potential of log files in order to address several questions about student performance in the knowledge acquisition phase of CPS (Greiff et al., 2015; Greiff, Molnár, Martin, Zimmermann, & Csapó, 2018; Greiff, Niepel, Scherer, & Martin, 2016; Molnár & Csapó, 2018). However, so far, the findings have been exclusively based on a student's score in the knowledge acquisition phase of CPS. This circumstance again highlights the necessity to extend our knowledge regarding the transition between knowledge acquisition and knowledge application in CPS, which will be done in the present study.

1.3 The Present Study

Taken together, the present study builds on several key findings of previous research. First, contemporary education aims at equipping students with the tools to successfully acquire and apply knowledge in increasingly complex contexts, which have proven to be two separate distinct processes, and CPS is a relevant showcase to investigate this transition of knowledge. Second, the two phases of knowledge acquisition and knowledge application are inherent in the process of CPS, and can be distinguished in the assessment of CPS performance. Third, earlier

studies have demonstrated the role of mental models in the knowledge acquisition phase of CPS, thereby making mental models a promising candidate for the differentiation between success and failure in the CPS knowledge transition of students.

However, at the same time we have illustrated several existing research gaps that we wish to address by means of this study. To begin with, up to now, the large amount of studies dealing with CPS has almost exclusively focused on how students can become better at acquiring knowledge. Hence, research on the transition between acquiring and applying knowledge in CPS is scarce at best. In addition, while significant latent correlations between the two CPS phases have been reported by previous research, ours will be a pioneering study in providing an analysis across the two phases on the individual student level. Moreover, while the phenomenon of being *'lost in transition'* (i.e., being unable to apply previously acquired knowledge successfully) has been present in the educational context for several decades, we are still unsure about the actual magnitude of this phenomenon (i.e., how large is the proportion of students affected by it?) in CPS. Therefore, the goal of our study is to analyze the following two research questions (RQs):

1. How many students who successfully solved the knowledge acquisition phase subsequently fail in the knowledge application phase (i.e., are 'lost in transition')?
2. Why do students who successfully solve the knowledge acquisition phase fail to successfully solve the knowledge application phase? Some specific hypotheses will be discussed individually, i.e.,:
 - a) Students generally fail to transfer their mental model from the knowledge acquisition over to the knowledge application phase

- b) Students generally fail to transfer their mental model efficiently from the knowledge acquisition over to the knowledge application phase
- c) Students are able to retrieve their mental model correctly in simple items, but fail to retrieve their model correctly in more complex items

2. Materials and Methods

2.1 Sample Characteristics

We analyzed the log files of a large-scale dataset of 9th grade students in Finland¹. Their age ranged from 13 to 18 years ($M = 15.70$, $SD = 0.44$). Of all $N = 1,369$ respondents, 701 were female and 668 were male. Written informed consent was required from both the students as well as their parents before participation. After processing the files according to our inclusion and exclusion criteria (see paragraph 2.5 Filtering and Statistical Analysis for further information), the final sample size used for statistical analysis was $N = 1,151$.

2.2 Measure

For the purpose of assessing students' CPS performance, we used the computer-based MicroDYN approach (Greiff et al., 2012), which is based on an underlying linear structural equation (LSE) framework that was formally introduced by Funke (2001).

Multiple independent complex problem environments are available in MicroDYN (see Figure 1 below for an example), each of which takes about five minutes to complete, when

¹ Please note that the dataset the findings of the present study are based on has been used in earlier publications (e.g., Krkovic, Greiff, Kupiainen, Vainikainen, & Hautamäki, 2014; Wüstenberg, Stadler, Hautamäki, & Greiff, 2014; Stadler, Niepel, & Greiff, 2016; see also Vainikainen, 2014 for an overview of the assessment battery). However, no existing publication has used the data in order to investigate the transition of students between the knowledge acquisition and the knowledge application phase. Thus, the research questions addressed and results being reported in the present study are entirely unique to the present study.

incorporating both the knowledge acquisition as well as the knowledge application phase. While each MicroDYN item employs a different cover story and variable names used to ensure that participants are unable to solve it based on prior knowledge in a given domain, all of the items share the same underlying framework. Hence, they differ only in terms of the number variables present, the number of existing relations between the variables, and with regard to the presence or absence of an ‘eigendynamic’, and can therefore be distinguished according to complexity based on these individual characteristics. MicroDYN items can either be of low (two existing variable relations and no autonomous change in variables, i.e., ‘eigendynamics’), medium (more than two existing variable relations, no ‘eigendynamic’; see Figure 1 below), or high complexity (more than two existing variable relations and ‘eigendynamics’; Stadler, Niepel, & Greiff, 2016). In addition, the validity of the MicroDYN assessment tool has been verified by previous research, particularly in relation to other CPS measurement approaches (e.g., Greiff et al., 2015).

Upon entering the first phase, knowledge acquisition (depicted on the left side of Figure 1), the participant is able to explore the system without any constraints for 180 seconds. During this period, they are allowed to make as many input variable manipulations (‘Blue Chips’, ‘Green Chips’, and ‘Red Chips’, see left part of the left side of Figure 1) in order to discover any effects on the output variables (‘Cards’, ‘Pawns’, and ‘Score’) as desired. At any time during this phase, the solver can draw arrows to indicate the connections between the respective variables (see bottom part of the left side of Figure 1), which serve as the key performance indicator in this first phase of CPS (see also paragraph 2.4 Scoring). Importantly, no arrows are shown to the participants at the beginning of the knowledge acquisition phase (i.e., the model is completely empty), and they are free to draw arrows between any input and output variables they personally

deem appropriate. At this point, the correct solution is not yet available to the solver, and they finish this phase without receiving feedback on their performance by simply clicking on “Done”.

Subsequently, the participant directly enters the second phase, knowledge application (shown on the right side of Figure 1). Now, specific target goals for the output variables are being presented (see right part of the right side of Figure 1), which should be reached in no more than four steps by, again, freely manipulating the input variables. In this phase, achieving these target goals represents the key indicator of performance (see also paragraph 2.4 Scoring). During the entire knowledge application phase, all existing variable relations are being shown to the participant (i.e., the *correct* model, irrespective of which arrows were drawn by the solver in the previous phase; see bottom part of the right side of Figure 1).

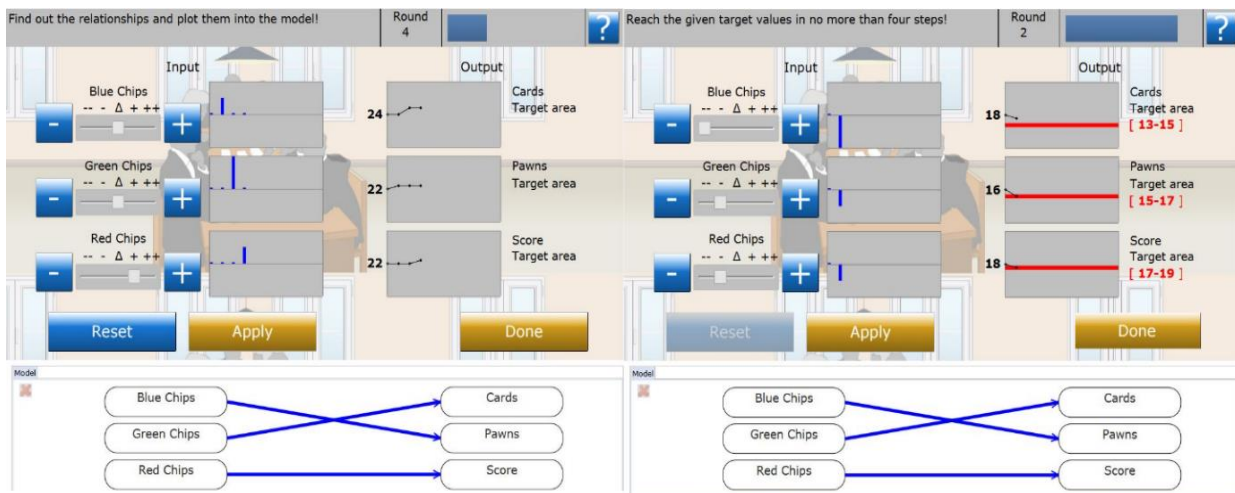


Figure 1. Screenshot from the knowledge acquisition phase (left part) and the knowledge application phase (right part) of the MicroDYN item ‘Game’, an item of medium complexity with three variable relations and without ‘eigendynamic’.

2.3 Procedure

In total, the participants solved nine MicroDYN items on individual school computers over a testing period of approx. 45 minutes. The individual items were presented to the students in a fixed, predefined order. Firstly, the participants received an instruction video about how to interact with MicroDYN. Afterwards, they were asked to solve the first five items (three of them being of low and two of medium complexity), followed by an instruction video for ‘eigendynamics’ (i.e., the ability of variables to change autonomously). Then, the participants received four additional items to solve (one of which being of medium, and three of high complexity).

2.4 Scoring

For each respective phase of a given MicroDYN item, the participants received a score based on their performance. In the knowledge acquisition phase, a participant who managed to draw all arrows indicating the existing relations between input and output variables correctly received a score of ‘1’ for success. If one or more arrows were missing, or in case the participant drew incorrect arrows, he or she received a ‘0’ indicating failure. For the subsequent knowledge application phase, a participant was assigned to one of the following categories, depending on her or his performance: Category 1 (no target reached, no target approximation), Category 2 (no target reached, target approximation), Category 3 (one target reached, no approximation on the second target), Category 4 (one target reached, approximation on the second target), or Category 5 (all targets reached; i.e., overall success). Approximation was coded as reduction in the distance of the achieved value of an output variable from the target goal value of that particular output variable, compared to its initial value at the beginning of the knowledge application phase.

The nine MicroDYN items the students worked on can be grouped according to their respective complexity following the argumentation by Stadler and colleagues (2016), which considered both the number of relations and the presence of ‘eigendynamics’ in a system as determinants of item complexity. Thus, the three items, which contain only two variable relations (i.e., ‘Lemonade’, ‘Drawing’, and ‘Cat’), can be considered as relatively easy items. In addition, the three items with more than two variable relations but without ‘eigendynamic’ (i.e., ‘Moped’, ‘Game’, and ‘Handball’), can be assigned to the category of medium complexity. Finally, the three items in which both more than two variable relations and an ‘eigendynamic’ is present (i.e., ‘Gardening’, ‘Spaceship’, and ‘Aid’), can be regarded as most complex items.

2.5 Filtering and Grouping

Firstly, we removed all participants from our dataset who did not interact with the system in the knowledge application phase (i.e., those who did not perform a single variable manipulation until the second phase was terminated) for all respective items. Therefore, $N = 218$ participants were excluded from the beginning, since they did try to successfully solve the knowledge application phase in any of the nine given items. This was done to prevent a possible classification of these participants in the ‘failure’ category for knowledge application.

Secondly, the participants can generally be grouped according to their performance in both respective phases (see Figure 2 below). More specifically, a participant who succeeds in the knowledge acquisition phase, can either also complete the knowledge application phase successfully (Group A), or fail in the knowledge application phase (Group B). In contrast, someone who fails in the knowledge acquisition phase can either succeed (Group C) or fail in the

knowledge application phase (Group D).

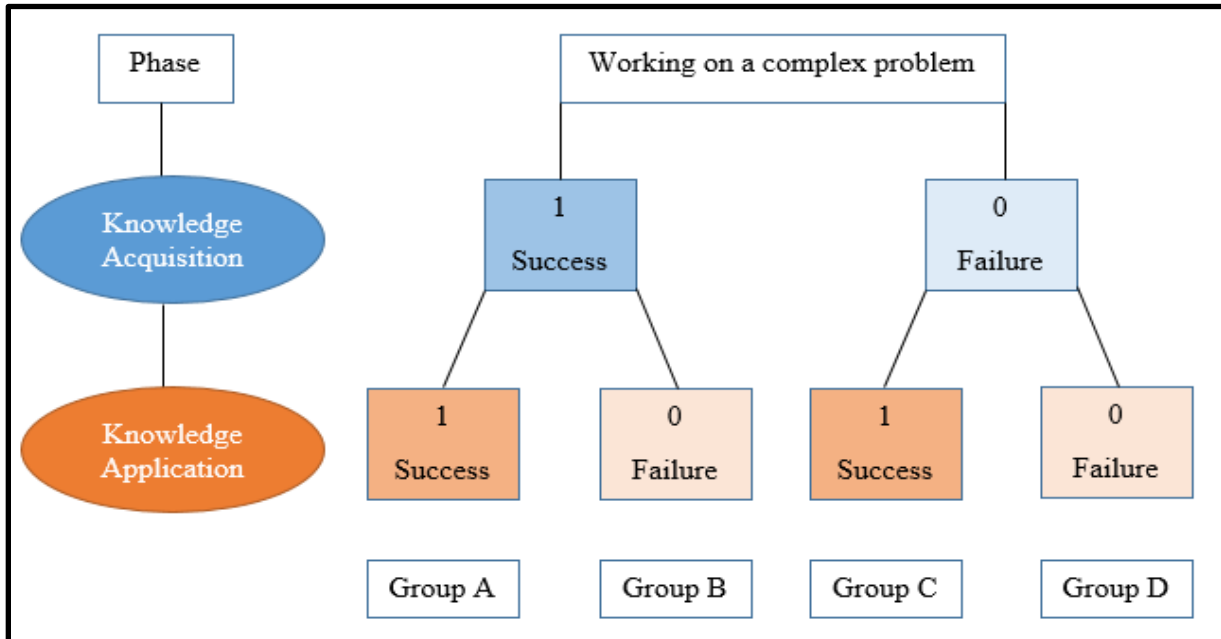


Figure 2. Decision Tree with all possible categories participants can be assigned to in a given MicroDYN item.

For our analyses, we were particularly interested in the students who succeeded in the knowledge acquisition phase but not knowledge application phase (i.e., Group B), in order to contrast these students to the ones who succeeded in both phases (i.e., Group A). Therefore, we created a grouping variable to include only the students in Groups A and B of each respective MicroDYN item in our subsequent analyses.

2.6 Statistical Analysis

Since the number of variable relations as well as the presence or absence of an ‘eigendynamic’ particularly contribute to the complexity of an item (Stadler et al., 2016), we

included these two factors in a Generalized Linear Mixed Model (GLMM; e.g., Baayen, Davidson, & Bates, 2008) analysis that was conducted in order to address the third hypothesis (i.e., part 'C') of RQ2. In order to address our question of how the complexity of an item influences the probability of a student being 'lost in transition', we firstly created a subset of the data selecting only the participants who succeeded in the knowledge acquisition phase of each given item for further analysis. This was done in order to ensure that for each item only those students who met the specifically relevant condition for being able to be 'lost in transition' in the first place were included in subsequent calculations, which is, to have mastered the knowledge acquisition phase successfully. Thereafter, we specified a GLMM incorporating the variables number of relations and 'eigendynamic' as fixed effects while simultaneously controlling for the variables participant and item in labelling them as random effects with varying intercepts. The dichotomous variable knowledge application score was used as dependent variable. Thus, our specified model can formally be defined as $Y_{ij} = \beta_o + b_{op} + b_{oi} + \beta_{1R} + \beta_{2E}$. In this equation, Y_{ij} denotes the logit of the probability of a successful knowledge application score. Additionally, β_o denotes the general intercept, while β_{1R} and β_{2E} reflect the two fixed item complexity characteristics number of relations and 'eigendynamic'. Lastly, the two varying intercepts for ability and item are controlled for in b_{op} and b_{oi} , respectively.

In order to perform the statistical analyses, the 25th version of the 'Statistical Package for the Social Sciences' (SPSS) software (IBM, 2017) was used for data transformation and grouping. We also used version 0.11.1 of JASP (JASP Team, 2019) for investigating RQs 1, 2a and 2b. In addition, version 4.0.0 of R (R Core Team, 2020) and the package lme4 (Bates, Maechler, Bolker, & Walker, 2015) were used for the GLMM analysis (RQ 2c).

3. Results

3.1 RQ1: How many students who solved the knowledge acquisition phase fail in the knowledge application phase?

Our first RQ addresses the question to what extent the problem of being ‘lost in transition’ is generally occurring. Therefore, we analyzed the relative frequencies of students who, after successfully solving the knowledge acquisition phase, did not manage to perform successfully in the knowledge application phase. The results of this analysis can be seen in Figure 3 below, indicating that, across all nine MicroDYN items, a significant proportion of students fell into this group (previously labelled as ‘Group B’, see Figure 2 above). More specifically, on average across all items, 42.05% of students who succeeded in the first, failed in the second phase. In addition, when grouping the percentages according to item complexity, the phenomenon affected on average 23.68% of students in easy items, 70.43% of students in items with medium complexity, and 32.02% of students in the most complex items with ‘eigendynamics’, respectively.

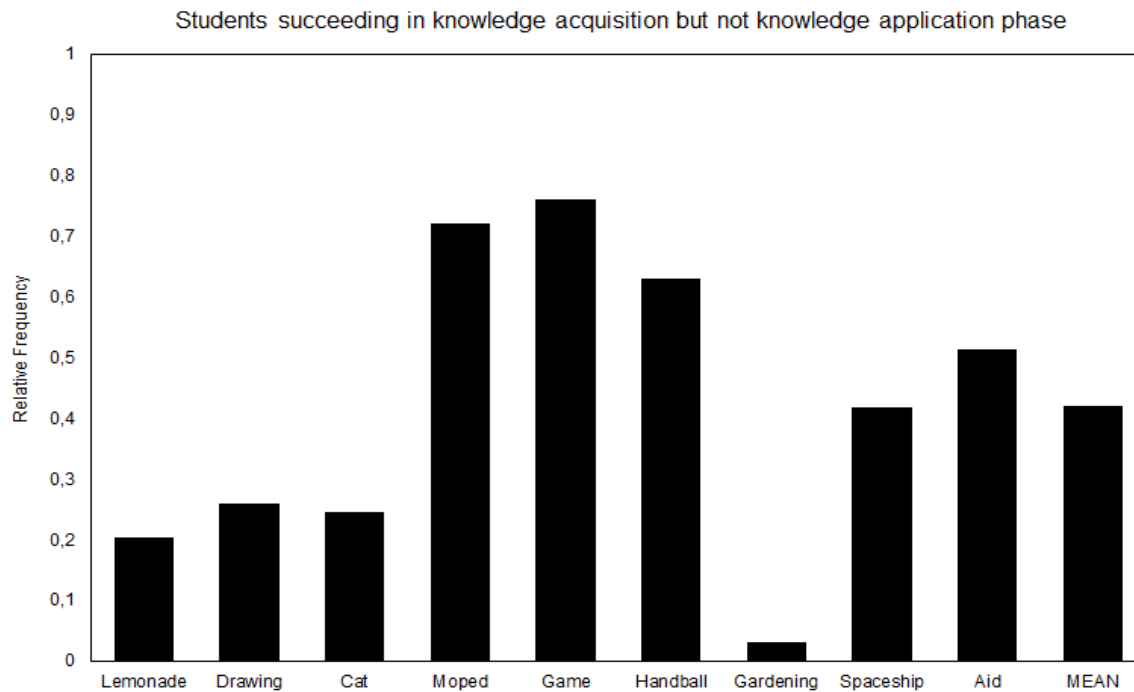


Figure 3. Relative frequency of students who are successful in the first (knowledge acquisition) but fail in the second (knowledge application) phase of CPS, for each respective MicroDYN item completed.

3.2 RQ2: Why Do Students Who Successfully Solve the Knowledge Acquisition Phase Fail to Successfully Solve the Knowledge Application Phase? – The Evaluation of Three Hypotheses

3.2.1 Hypothesis RQ2a: Students generally fail to transfer their mental model from the knowledge acquisition over to the knowledge application phase

The first potential explanation for the observed phenomenon is that students generally are unable to transfer the mental model they constructed during the knowledge acquisition phase over to the knowledge application phase. This would mean that they not only are unable to reach all target goals, but also that they do not manage to approximate or reach parts of the given target goals. That is why we analyzed the relative frequencies of students in a more fine-grained

manner, by grouping them according to the five different categories consisting of participants who, in the knowledge application phase, 1) neither approximated nor reached a single target goal, 2) did not reach but approximated one target goal, 3) reached one target goal but did not approximate the additional target goal(s), 4) reached one target goal and approximated the additional target goal(s), and those who 5) reached all target goals (cf. paragraph 2.3 above).

For this hypothesis, the students in Category 1 are particularly relevant. The results for this category for each respective MicroDYN item can be seen in Figure 4 below (indicated by the red part of each item-based bar), indicating that, on average across all nine MicroDYN items, 4.47% of the students fell into Category 1. Hence, only a minority of students who failed in the application phase did not reach or at least *approximate* one goal. In other words, over 95% of students were able to transfer “some” knowledge from the acquisition phase over to the knowledge application phase. That is why we consider the possibility that students were *generally* unable to transfer their mental model as unlikely.

3.2.2 Hypothesis RQ2b: Students generally fail to transfer their mental model efficiently from the knowledge acquisition over to the knowledge application phase

Therefore, it is possible that students *generally* are able to transfer their mental model from the first over to the second phase of CPS, yet, that this transfer is not *efficient*. Under the given circumstances, we defined efficiency as the ability to reach the target goals of a respective MicroDYN item within the predefined four steps. Thus, while students might be unable to reach all target goals within this limited amount of steps, some indicators that they can transfer at least a partly correct mental model can be investigated. These indicators are that students either reach one of multiple target goals, or that they at least approximate one of the target goals.

In order to analyze this hypothesis, we also relied on the grouping of the students mentioned above, this time with a particular emphasis on the students in Categories 2, 3 and 4, respectively (depicted in the orange, yellow, and light green bars of each completed item in Figure 4 below). Results showed that, on average across all nine items, 20.35% of students were part of Category 2, while 1.64% of participants fell in Category 3, and 15.77% were classified as being in Category 4; thus, arriving at a cumulated percentage of 37.76% of students who fell in one of these three categories. As also previously concluded in RQ2a, these outcomes indicate that some knowledge has been transferred from the knowledge acquisition to the application phase, but that students were unable to apply their knowledge efficiently to solve an item in a predefined, restricted number of steps. Now that we have, as indicated throughout the results section, obtained evidence in favor of this hypothesis, the question why students are unable to transfer their mental model efficiently from the first to the second phase leads to the possibility of item complexity playing a role, which will be discussed in the subsequent hypothesis below.

3.2.3 Hypothesis RQ2c: Students are able to retrieve their mental model correctly in simple items, but fail to retrieve their mental model correctly in more complex items

Since we know that the MicroDYN items used in this study can be distinguished according to their complexity by means of the number of existing variable relations and the presence or absence of an ‘eigendynamic’, we hypothesized that these factors might play a key role for the accuracy of the mental model transfer of students. In other words, we were expecting that determinants of item complexity (i.e., number of relations and presence of ‘eigendynamic’) play a critical role in transferring knowledge from the acquisition phase to the application phase.

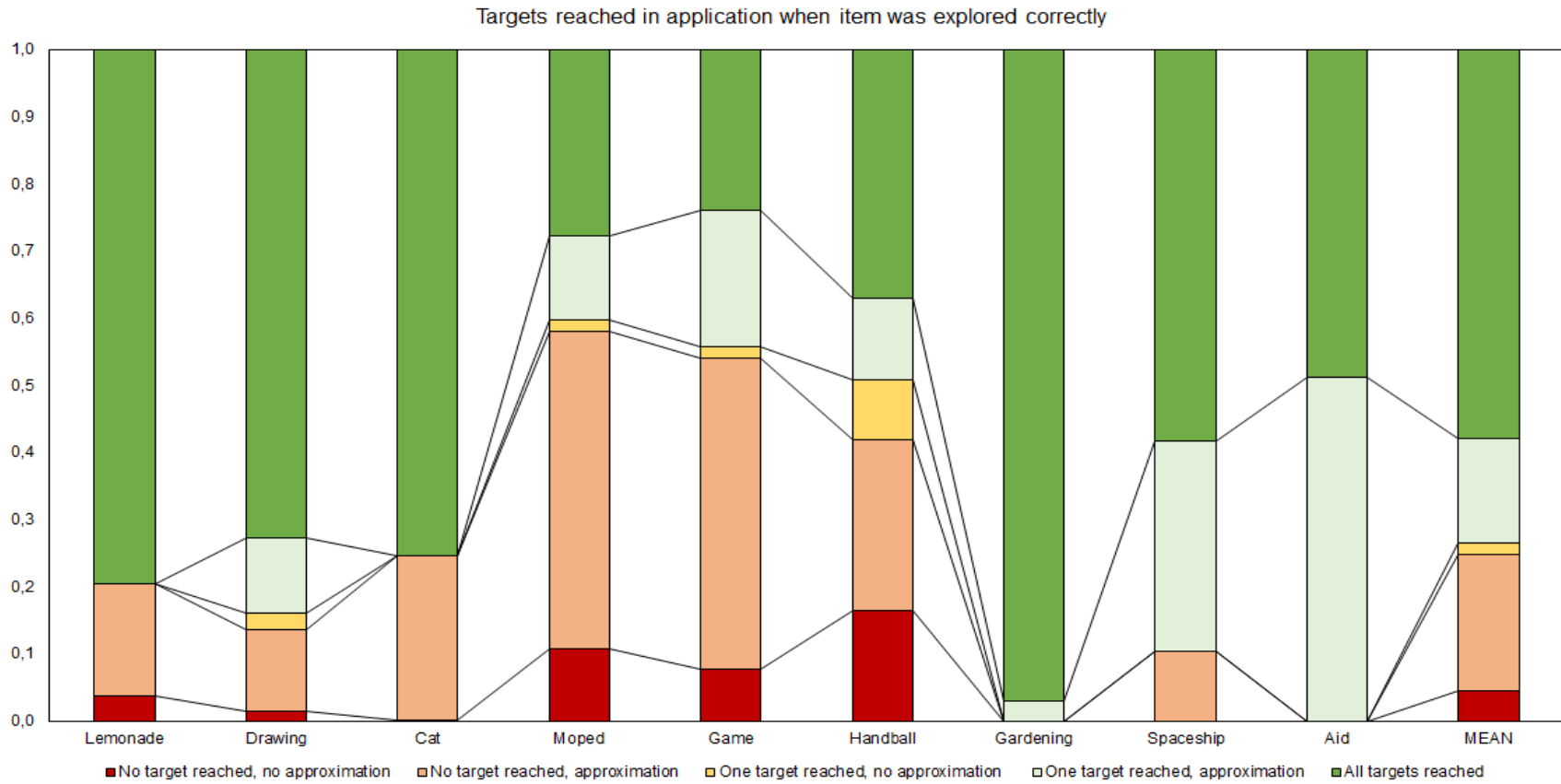


Figure 4. Relative frequencies of students in the respective five categories according to their performance in the knowledge application phase of CPS across all nine MicroDYN items.

In order to test this hypothesis, our Generalized Linear Mixed Model (GLMM) included number of relations and ‘eigendynamic’ as fixed effects, and participant and item as effects with random intercepts. With this procedure, we were able to test whether the number of relations or the presence of ‘eigendynamics’ exerted a significant influence on participants solving the acquisition but not the application phase while simultaneously controlling for participant and item. Overall, results of the GLMM analysis showed that while the presence or absence of an ‘eigendynamic’ did not appear to be statistically significant ($\beta_{2E} = .58, p = .11$), the number of relations present in an item significantly predicted ($\beta_{1R} = -2.44, p < .001$) how a student who completed the first phase of CPS successfully, scored in the subsequent second phase of CPS. In other words, the more relations between variables are present in an item, the less likely are students who completed the knowledge acquisition phase successfully to also be successful in the knowledge application phase. The results are summarized in Table 1 below. We can conclude that these outcomes support our hypothesis H2c, since one of the two complexity predictors proved to exert significant influence on our dependent variable. More specifically, in items with fewer variable relations, participants were more likely to succeed in the knowledge application phase.

Table 1. *Results of the Generalized Linear Mixed Model (GLMM) investigating the effects of the number of relations and ‘eigendynamic’ for the score in the knowledge application phase when controlling for participant and item.*

Score Knowledge Application Phase				
Fixed Effects	β	<i>SE</i>	<i>Z</i>	<i>p</i>
Number of relations	-2.44	0.40	-4.13	<0.001***
‘Eigendynamic’	0.58	0.67	1.62	0.11
Random Effects Intercept	Variance	<i>SD</i>		
Participant	2.34	1.53		
Item	0.67	0.82		
Marginal R ²	0.19			
Conditional R ²	0.58			

4. Discussion

The overarching aim of this study was twofold. To begin with, we wanted to uncover the magnitude of what we call the ‘lost in transition’ phenomenon, which refers to the case that a participant successfully explores the hidden variable relations in the knowledge acquisition phase, but fails in reaching the predefined target goals in the following knowledge application

phase (i.e., Group B as shown in Figure 2), in the particular showcase of CPS. In addition, we wanted to investigate potential reasons for why such a large group of students is affected by this phenomenon. We will now discuss the results with regard to their implications, particularly in light of their relevance for the factors to be taken into account by upcoming training programs of CPS.

4.1 The Magnitude of the ‘Lost in Transition’ Phenomenon in Complex Problem Solving

Our initial analysis revealed that, in our large-scale dataset, on average almost one in two students (i.e., 42.05%) was unable to reach all target goals in the second phase of CPS, after she or he had initially explored all variable connections successfully. Given the generally high latent correlations between the two phases of CPS (e.g., Herde et al., 2016), it comes as a surprise that so many students experience this situation.

Still, we wish to acknowledge that differences in percentages of students affected were present when distinguishing item complexity. In easy items, fewer students fell into this Group B, namely only about one in four (i.e., 23.68%). However, as soon as item complexity augmented, so did the percentages of students in this group. For items with more variable relations but without ‘eigendynamics’ (medium complexity), 70.43% of students were affected. When it comes to the most complex items with a simultaneously high number of variable relations and ‘eigendynamics’, about one in three students (i.e., 32.02%) was affected. One particular outlier, as can be seen in Figure 3, was found in the ‘Gardening’ item, which was the first item with ‘eigendynamic’ the students were confronted with. Therefore, while the comparatively low percentage of students being ‘lost in transition’ for this item (i.e., 3.03%) may come unexpectedly, it can be traced back to the fact that very few students solved the knowledge

acquisition phase successfully in the first place, namely only 66 (i.e., 6.76% of all) participants. For the two subsequent items containing ‘eigendynamics’, 41.79% and 51.24% of students experienced the phenomenon, respectively.

Overall, we are able to recognize a trend that over the course of increasing item complexity, more students tend to fall into Group B, while the rates drop slightly when the students attempt the items with ‘eigendynamics’. This might happen due to the students already having familiarized themselves with the way MicroDYN assesses their CPS skills by the time these types of items are introduced. Given these generally high rates of students experiencing success in phase one and failure in phase two of CPS, we have identified a formerly unknown performance gap in CPS, which requires further investigation in order to find potential reasons for its occurrence.

4.2 Possible Explanations Why the ‘Lost in Transition’ Phenomenon Occurs

We analyzed three hypotheses (RQs2a, b, and c) for the occurrence of the aforementioned performance gap between the two phases of CPS. Each of these hypotheses will be evaluated individually below, before discussing the overarching implications of our findings.

Firstly, we hypothesized that the students in Group B are generally unable to transfer the mental model they created successfully in the knowledge acquisition phase, over to the knowledge application phase. As already indicated by the results, although there is a proportion of students who did not manage to approximate a single target goal, this proportion was rather small, with an average of 4.47% of students across all nine items. Taken together, while we can acknowledge that some students may have been unable to translate their mental model *in*

general, it is by far not the majority. Therefore, we are able to conclude that this explanation is an unlikely underlying driving force for the performance gap between the two phases of CPS.

Secondly, we investigated the possibility that the participants are generally able to transfer their mental model from the first to the second phase of CPS; however, that this transfer is not carried out with sufficient efficiency in order to reach all target goals. As we can infer based on the results, a largely higher proportion of students (37.76% on average across all nine items) was able to reach or at least approximate one of the target goals, indicating that a transfer of the mental model was present to some extent. Thus, it can be concluded that the efficiency of the mental model transfer represents a key factor for the performance gap between knowledge acquisition and knowledge application. In order to find out more about the circumstances under which students are more or less efficient in transferring their mental model, we investigated the following third hypothesis.

Thirdly, based on previous research, we hypothesized that the efficiency of the mental model transfer depends on the complexity of a given MicroDYN item (Stadler et al., 2016). With regard to item complexity, we analyzed whether the number of relations and/or the presence of an ‘eigendynamic’ in an item had significant impact on a student experiencing the ‘lost in transition’ phenomenon. The results with regard to the third hypothesis carry several implications. Firstly, they show that the more relations were present in an item, the higher the probability that a student who successfully completed the knowledge acquisition phase did not succeed in the subsequent knowledge application phase. Secondly, the presence of ‘eigendynamics’ did not produce a significant effect. On the one hand, this finding represents a contrast to earlier research that has found ‘eigendynamics’ to play a significant role for CPS

success (e.g., Stadler et al., 2016). On the other hand, these findings are exclusively based on success in the knowledge acquisition phase. Given the different requirements of the two distinct phases, our results are in line with previous studies that were unable to identify the presence of ‘eigendynamics’ as important predictor of item difficulty in the knowledge application phase (Greiff, Krkovic, & Nagy, 2014).

Overall, the two predictors explained 19% of the variance in our dependent variable, while the full GLMM was able to explain 58% of the variance in the knowledge application score when also taking participant and item as random effects into account. Therefore, we can conclude that the efficiency of a student’s mental model transfer from the first to the second phase of CPS largely depends on the number of relations that are present in an item; thus, supporting this hypothesis. In items with fewer relations, students are more likely to be able to efficiently transfer their mental model, and to reach all target goals.

4.3 Main Implications With Regard to Learning Analytics

Several main implications can be inferred from our results. First and foremost, despite the generally high latent correlations between the knowledge acquisition and the knowledge application phase of CPS obtained in previous studies, the previously unknown ‘lost in transition’ phenomenon affects a great proportion of students. Therefore, it is crucial to openly communicate the existence of this problem in order to adapt future CPS training programs accordingly, which might also help students for their overarching educational success.

When carefully considering the requirements of each of the two respective CPS phases, however, our findings do not come as a great surprise. The focus of the knowledge acquisition phase lies in the almost unrestrained discovery of the relations between the input and output

variables by means of systematic variation, sometimes also referred to manipulating one input variable at a time, or the 'VOTAT' strategy (e.g., Schwichow, Croker, Zimmermann, Höffler, & Härtig, 2016; Tschirgi, 1980). Subsequently, the requirements drastically shift to the goal-directed manipulation of the input variables within only four attempts that can be made. Thus, allowing a participant to make only four variable manipulations and simultaneously imposing multiple target goals that have to be reached, renders his or her carefully constructed methodology of working on such a microworld impossible to use in this second phase of CPS. For instance, the previously useful VOTAT strategy is suddenly obstructive instead of beneficial for success, and the student needs to recognize the impact of such a rigorous shift in demands for his or her *modus operandi* immediately. However, if the participant takes one or two variable manipulations before noticing this shift, it might already be too late to reach the target goals. Therefore, the cognitive flexibility of students to immediately adapt their mental model to this new situation and understand the implications of the changed task demands before interacting with the system, may determine their ability to succeed in the knowledge application phase (Canas, Quesada, Antolí, & Fajardo, 2003; Kreams, 1995). As such, metacognitive aspects such as planning, monitoring and reflecting of and on how a complex problem should be/is being tackled should become a hallmark of future CPS training programs (McLoughlin & Hollingworth, 2002; Rudolph, Niepel, Greiff, Goldhammer, & Kröner, 2017), instead of only focusing on teaching students a particular strategy such as VOTAT (e.g., Wüstenberg, Stadler, Hautamäki, & Greiff, 2014).

More generally, when investigating the way people learn new and complex skills, Anderson (1982) proposes a framework, which distinguishes between declarative and procedural knowledge, arguing for the necessity of breaking down its individual steps in order to be able to

transfer knowledge from declaration to procedure successfully when trying to learn a new cognitive skill. Such a ‘knowledge compilation’ (p. 2) process closely resembles the aforementioned metacognitive aspects of monitoring and reflecting in CPS (Gott, Lajoie, & Lesgold, 1991; Rudolph et al., 2017). Similarly, the importance of such a metacognitive aspects in reflecting on previous and refining one’s future steps recently has been shown to play an important role when engaging in problem solving (Anderson & Fincham, 2014), game-based learning (Taub, Azevedo, Bradbury, & Mudrick, 2020), and in video gaming research (Anderson, Betts, Bothell, Hope, & Lebiere, 2019). For CPS, this may translate to the crucial aspect of a student being able not only to show that they know how given variables are connected, but also to use this knowledge to achieve certain goals. This way, the relevance for such metacognitive skills to be addressed in future CPS training programs is highlighted once more.

While previous research has already intended to train the CPS competence in students, results have often failed to show the desired outcomes of helping students to improve their performance in knowledge acquisition and knowledge application simultaneously. For instance, in a CPS training study by Kretzschmar and Süß (2015), participants were assigned to a ‘learning-by-doing’ training program, in which they were confronted with several different CPS microworlds, to improve their CPS performance. Results showed that, while performance improved with regard to exploration of a given system, the training did not have any beneficial effect on the students being able to apply the previously gained knowledge in order to reach the given target goals. Similarly, previous research has revealed that knowledge about a particular strategy is an insufficient predictor for actual CPS performance (e.g., Wüstenberg et al., 2014).

Overall, the results of our study point towards the necessity for future CPS training programs to take metacognitive factors into account, since these factors apparently play a predominant role for success in both phases of CPS. However, how can this be done properly in order to prevent the ‘lost in transition’ phenomenon? Some suggestions have been made in previous studies. For instance, directly instructing students in how to work on a specific complex problem has shown to outperform allowing students to freely explore a system while providing only marginal guidance (Kluge, 2008). However, as described by Paas and Van Merriënboer (1994), it is important to also consider the participants’ cognitive constraints. Two potentially useful approaches that have been successful in the past have used similar mechanisms. The first approach employs a hierarchical model, which differentiates the existing goals and orders them according to their importance for a given task, and then continues to provide suitable strategies to reach each goal, and lower-order competencies eliciting these strategies (Frederiksen & White, 1989). This approach leads to separate task components being trained before students are confronted with the entire task. With regard to CPS, students could be provided with or asked to generate such a goal hierarchy for each respective phase, in order to be able to immediately know what is expected of them at any moment during a CPS task. The second approach by Gopher, Weil, and Siegel (1989) is termed emphasis-manipulation approach and focuses on different subcomponents of a given task, which are trained separately after one another, leading to greater flexibility in how a task can be completed (e.g., by adaptively identifying the subcomponent that is relevant at a given point in time). This could be achieved in CPS for instance by initially having participants working on the knowledge acquisition exclusively, in order to understand which strategies are beneficial. After having mastered this task, participants would then be trained for the knowledge application phase accordingly, before having to work

on an item including both phases. In addition, more recent approaches have emphasized the benefit of thinking aloud protocols (e.g., Barrett et al., 2013), and of deliberately planning which variable manipulation to perform next (Eichmann, Goldhammer, Greiff, Pucite, & Naumann, 2019). Ideally, several of these methodological approaches to provide scaffolding for students could be combined in a training program, for example, by providing them with individual trainings pieces targeted at specific strategies for each respective phase as well as how to properly plan, monitor, and evaluate how one approaches a complex problem.

In sum, many different ways have been proposed to foster metacognitive aspects in CPS, which have been shown to be relevant not only for improving success in CPS performance in general, but also seem to be a promising candidate for avoiding the ‘lost in translation’ phenomenon in the future. In addition, future CPS training programs that are more fine-grained and consider multiple key elements of the CPS process will allow students to become better complex problem solvers, which, in turn, will increase the chances of them succeeding along their educational paths.

However, the described phenomenon is not exclusively linked to learning analytics in CPS. In a broader context, transferring learned knowledge from one phase or step to another is an integral part of nearly every learning situation including transversal competencies, such as collaborative problem solving (Graesser et al., 2018; Herborn, Stadler, Mustafić, & Greiff, 2018), scientific inquiry (Chen & Klahr, 1999), or dynamic decision making (Karakul & Qudrat-Ullah, 2008). The aim of this study was to shed light on the ‘lost in transition’ phenomenon in a specific context, in which the two phases knowledge acquisition and knowledge application are clearly separated and, therefore, the phenomenon can be observed under “clean” conditions.

Future studies now need to, firstly, further investigate why this phenomenon occurs, by either leveraging the potential of even more fine-grained process data as in the present study (Greiff et al., 2018; Molnár & Csapó, 2018; Stadler, Fischer, & Greiff, 2019), or by combining correlational and experimental approaches in order to determine the cognitive factors underlying the phenomenon (e.g., Krieger, Zimmer, Greiff, Spinath, & Becker, 2019). Secondly, future research should try to generalize this phenomenon to other tasks and domains, and, thirdly, build and adapt cognitive models on how transferring knowledge should be assessed, addressed, and taught in order to ensure efficient learning for students, and get closer to filling the crucial performance gap between knowledge acquisition and knowledge application in educational contexts.

4.4 Limitations

Some limitations of the present study have to be acknowledged. Firstly, our results are based on a single dataset of students from a single country. However, the students worked on multiple instead of only one CPS item, and the sample was carefully selected and is representative of the Finnish student population regarding its demographic and socioeconomic characteristics (Vainikainen, 2014). Still, investigations with different datasets provide an interesting avenue for future research to obtain more information on the ‘lost in transition’ phenomenon, and its magnitude and relevance for students in different countries and at different ages.

Secondly, the students only worked on items from a single type of microworld, namely MicroDYN. Thus, the question arises if our results are generalizable also to other types of CPS assessments that clearly distinguish between the two phases of knowledge acquisition and

knowledge application, such as Space Shuttle (Wirth & Funke, 2005), the Genetics Lab (Sonnleitner et al., 2013), or Multiflux (Bühner et al., 2008). As already outlined above, future studies should therefore aim at replicating our analyses with data from different CPS assessment tools but also from other domains, in order to cross validate the existence of the ‘lost in transition’ phenomenon.

Thirdly, while we believe that the MicroDYN framework fitted well for the purpose of our study, its inherent clear separation of the two CPS phases represents a feature with limited transferability to problem solving in the real world. In real world problem solving, people usually acquire and apply knowledge simultaneously, and also more frequently alternate between these two phases (e.g., Sarathy, 2018). Therefore, the investigation of the ‘lost in transition’ phenomenon should be extended to assessment tools in which the two phases are more intertwined as it occurs in natural settings. However, we believe to have established a pertinent showcase of the ‘lost in transition’ phenomenon by means of evaluating it with the current approach, since it remains unclear how these two phases can be reliably separated in real life settings.

5. Conclusion

The primary aim of this study was to uncover the magnitude of the ‘lost in transition’ phenomenon, referring to the case that a participant who successfully acquires knowledge fails to subsequently apply this knowledge, in CPS. In addition, we discussed some potential reasons for why this phenomenon arises. Overall, we can conclude that a significant amount of students experiences being ‘lost in transition’, and that the probability of this phenomenon occurring particularly depends on the number of variable relations present in an item. Moreover, most

students are able to at least approximate or reach one of the target goals, indicating that, instead of a general mental model transfer error being the primary cause, it is rather the absent efficiency of this transfer representing the driving force for the occurrence of this phenomenon.

In closing, we have provided a new and important viewpoint to the process and research of CPS in being the first study analyzing the transition between its two phases on an individual instead of a latent level. Several implications are targeted at the role of CPS in light of learning analytics. In particular, we have paved the way for several underlying facets of the overarching CPS process to be included in future educational training programs, by making explicit implementation suggestions. Primarily, these include equipping students with the strategic and metacognitive tools (e.g., cognitive flexibility, planning, monitoring, and evaluating) for performing successfully in each distinct phase, thereby minimizing the possibility to experience being ‘lost in transition’ in the future. In addition, our findings aid the educational endeavor to make students better overall complex problem solvers, which simultaneously carries several potential benefits for their scholastic performance, and represents a suitable means to help them overcome the performance gap between knowledge acquisition and knowledge application.

References

- Adams, R., Vista, A., Scoular, C., Awwal, N., Griffin, P., & Care, E. (2015). Automatic coding procedures for collaborative problem solving. In *Assessment and teaching of 21st century skills* (pp. 115-132). Springer, Dordrecht.
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, 89(4), 369.
- Anderson, J. R., Betts, S., Bothell, D., Hope, R., & Lebiere, C. (2019). Learning rapid and precise skills. *Psychological review*, 126(5), 727.
- Anderson, J. R., & Fincham, J. M. (2014). Extending problem-solving procedures through reflection. *Cognitive psychology*, 74, 1-34.
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4), 1279-1333. <https://doi.org/10.1162/003355303322552801>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Barrett, J. D., Peterson, D. R., Hester, K. S., Robledo, I. C., Day, E. A., Hougen, D. P., & Mumford, M. D. (2013). Thinking About Applications: Effects on Mental Models and Creative Problem-Solving. *Creativity Research Journal*, 25, 199–212. <https://doi.org/10.1080/10400419.2013.783758>
- Bates, D., Maechler, M., Bolker, B. and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01

- Buchner, A. (1995). Basic topics and approaches to the study of complex problem solving. In: P.A. Frensch & J. Funke (Eds.). *Complex problem solving: the European perspective*. Hillsdale, NJ: Lawrence Erlbaum.
- Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual–spatial-intelligence and their relationship to problem-solving. *Intelligence*, 36(6), 672-680.
- Canas, J., Quesada, J., Antolí, A., & Fajardo, I. (2003). Cognitive flexibility and adaptability to environmental changes in dynamic complex problem-solving tasks. *Ergonomics*, 46(5), 482-501.
- Charland, P., Léger, P. M., Cronan, T. P., & Robert, J. (2016). Developing and assessing ERP competencies: Basic and complex knowledge. *Journal of Computer Information Systems*, 56(1), 31-39.
- Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child development*, 70(5), 1098-1120.
- Darabi, A. A., Nelson, D. W., & Seel, N. M. (2009). Progression of mental models throughout the phases of a computer-based instructional simulation: supportive information, practice, and performance. *Computers in Human Behavior*, 25, 723–730.
doi: 10.1016/j.chb.2009.01.009
- Dörner, D., & Funke, J. (2017). Complex problem solving: *What it is and what it is not*. *Frontiers in Psychology*, 8, 1–11. <https://doi.org/10.3389/fpsyg.2017.01153>
- Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., & Naumann, J. (2019). The role of planning in complex problem solving. *Computers & Education*, 128, 1-12.
<https://doi.org/10.1016/j.compedu.2018.08.004>

- Everwijn, S., Bomers, G., & Knubben, J. (1993). Ability- or competence-based education: bridging the gap between knowledge acquisition and ability to apply. *Higher Education*, 25, 425-38. <https://doi.org/10.1007/BF01383845>
- Frederiksen, J. R., & White, B. Y. (1989). An approach to training based upon principled task decomposition. *Acta Psychologica*, 71(1-3), 89-146.
- Funke, J. (1993). Microworlds based on linear equation systems: A new approach to complex problem solving and experimental results. In *The cognitive psychology of knowledge* (pp. 313-330). Elsevier Science Publishers.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & reasoning*, 7(1), 69-89. <https://doi.org/10.1080/13546780042000046>
- Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 682-685). Heidelberg: Springer.
- Gašević, D. (2019). Using learning analytics to measure 21st-century skills. In *Proceedings of the Research Conference of the Australian Council for Educational Research 2019*, 46-50.
- Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist*, 50(1), 43-57. <https://doi.org/10.1080/00461520.2014.999919>
- Gobert, J. D., Sao Pedro, M., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: Measuring students' science inquiry skills using educational data mining. *Journal of the Learning Sciences*, 22(4), 521-563. <https://doi.org/10.1080/10508406.2013.837391>

- Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica*, *71*(1-3), 147-177.
- Gott, S. P., Lajoie, S. P., & Lesgold, A. (1991). Problem solving in technical domains: How mental models and metacognition affect performance. In *Instruction: Theoretical and applied perspectives* (pp. 107-117). New York, NY, US: Praeger Publishers
- Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest*, *19*(2), 59-92. <https://doi.org/10.1177/1529100618808244>
- Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015). Assessing complex problem-solving skills with multiple complex systems. *Thinking & Reasoning*, *21*(3), 356-382. <https://doi.org/10.1080/13546783.2014.989263>
- Greiff, S., Krkovic, K., & Nagy, G. (2014). The systematic variation of task characteristics facilitates the understanding of task difficulty: A cognitive diagnostic modeling approach to complex problem solving. *Psychological Test and Assessment Modeling*, *56*(1), 83-103.
- Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education*, *126*, 248-263. <https://doi.org/10.1016/j.compedu.2018.07.013>
- Greiff, S., Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving: An analysis of behavioural

data from computer-generated log files. *Computers in Human Behavior*, *61*, 36–46.

<https://doi.org/10.1016/j.chb.2016.02.095>

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity of complex problem solving measures. *Intelligence*, *50*, 100-113.

<https://doi.org/10.1016/j.intell.2015.02.007>

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015). Computer-generated log file analyses as a window into students' minds?: A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education*, *91*, 92–105.

<https://doi.org/10.1016/j.compedu.2015.10.018>

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

<https://doi.org/10.1177/0146621612439620>

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

364–379. <https://doi.org/10.1037/a0031856>

Halasz, F. G., & Moran, T. P. (1983). Mental models and problem solving in using a calculator.

In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 212-216), ACM.

Harris, C. J., Krajcik, J., Pellegrino, J. & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational Measurement: Issues and*

Practice, *38*(2), 53-67.

- Herborn, K., Stadler, M., Mustafić, M., & Greiff, S. (in press). The assessment of collaborative problem solving in PISA 2015: Can computer agents replace humans?. *Computers in Human Behavior*, 2018. <https://doi.org/10.1016/j.chb.2018.07.035>
- 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. <https://doi.org/10.1080/08957347.2016.1209208>
- IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.
- JASP Team (2019). JASP (Version 0.11.1) [Computer software].
- Karakul, M., & Qudrat-Ullah, H. (2008). How to improve dynamic decision making? Practice and promise. In *Complex decision making* (pp. 3-24). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73665-3_1
- Kay, K., & Greenhill, V. (2011). Twenty-first century students need 21st century skills. In *Bringing schools into the 21st century* (pp. 41-65). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0268-4_3
- Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education*, 127, 233-251. <https://doi.org/10.1016/j.compedu.2018.08.023>

Kluge, A. (2008). What you train is what you get? Task requirements and training methods in complex problem-solving. *Computers in Human Behavior*, 24(2), 284-308.

<https://doi.org/10.1016/j.chb.2007.01.013>

Kolfschoten, G., Lukosch, S., Verbraeck, A., Valentin, E. & de Vreede, G.-J. (2010). Cognitive learning efficiency through the use of design patterns in teaching. *Computers &*

Education, 54, 652-660. <https://doi.org/10.1016/j.compedu.2009.09.028>

Krems, J. F. (1995). Cognitive flexibility and complex problem solving. In: P.A. Frensch & J.

Funke (Eds.). *Complex problem solving: the European perspective*. Hillsdale, NJ:

Lawrence Erlbaum.

Kretschmar, A., & Süß, H.-M. (2015). A Study on the Training of Complex Problem Solving Competence. *Journal of Dynamic Decision Making*, 1, 1-14.

DOI:10.11588/jddm.2015.1.15455

Krieger, F., Zimmer, H. D., Greiff, S., Spinath, F. M., & Becker, N. (2019). Why are difficult figural matrices hard to solve? The role of selective encoding and working memory

capacity. *Intelligence*, 72. <https://doi.org/10.1016/j.intell.2018.11.007>

Krkovic, K., Greiff, S., Kupiainen, S., Vainikainen, M.-P., & Hautamäki, J. (2014). Teacher evaluation of student ability. What roles do teacher gender, student gender, and their

interaction play? *Educational Research*, 56, 243-256.

<http://dx.doi.org/10.1080/00131881.2014.898909>

Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, 33(4), 347-368.

- Kurban, C. F. (2018). Designing effective, contemporary assessment on a flipped educational sciences course. *Interactive Learning Environments*, 27(8), 1143-1159.
<https://doi.org/10.1080/10494820.2018.1522650>
- Laurillard, D. (2012). *Teaching as a design science: Building pedagogical patterns for learning and technology*. New York, NY: Routledge.
- Lim, L. A., Gentili, S., Pardo, A., Kovanović, V., Whitelock-Wainwright, A., Gašević, D., & Dawson, S. (2019). What changes, and for whom? A study of the impact of learning analytics-based process feedback in a large course. *Learning and Instruction*.
<https://doi.org/10.1016/j.learninstruc.2019.04.003>
- Mainert, J., Kretzschmar, A., Neubert, J. C., & Greiff, S. (2015). Linking complex problem solving and general mental ability to career advancement: Does a transversal skill reveal incremental predictive validity?. *International Journal of Lifelong Education*, 34(4), 393-411. <https://doi.org/10.1080/02601370.2015.1060024>
- Mayer, R.E., & Wittrock, M. C. (2006). Problem solving. In P.A. Alexander & P.H. Winne (Eds.) *Handbook of educational psychology*, (2nd ed., pp. 287–304). Mahwah, NJ: Lawrence Erlbaum.
- McLoughlin, C., & Hollingworth, R. (2002). *Bridge over troubled water: Creating effective online support for the metacognitive aspects of problem solving* (pp. 1267-1272). Association for the Advancement of Computing in Education (AACE).
- Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in psychology*, 9, 302. <https://doi.org/10.3389/fpsyg.2018.00302>

- Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2014). Extending the assessment of complex problem solving to finite state automata. *European Journal of Psychological Assessment*, 31, 181-194. <https://doi.org/10.1027/1015-5759/a000224>
- Nistor, N., & Hernández-García, Á. (2018). What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior*, 89, 335-338. <https://doi.org/10.1016/j.chb.2018.07.038>
- Novick, L. R., & Bassok, M. (2005). Problem solving. *The Cambridge handbook of thinking and reasoning*, 137, 320-350.
- OECD. (2009). *PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science*. Paris, France: OECD Publishing.
- OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris, France: OECD Publishing. Retrieved from http://www.oecd-ilibrary.org/education/pisa-2012-assessment-and-analyticalframework_9789264190511-en
- OECD (2014). *PISA 2012 results: Creative problem solving: Students' skills in tackling real-life problems (Volume V)*. Paris, France: OECD Publishing.
- Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational psychology review*, 6(4), 351-371.
- R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Rickheit, G., & Habel, C. (1999). *Mental models in discourse processing and reasoning* (Vol. 128). Amsterdam, NL: Elsevier.

- Rudolph, J., Niepel, C., Greiff, S., Goldhammer, F., & Kröner, S. (2017). Metacognitive confidence judgments and their link to complex problem solving. *Intelligence*, *63*, 1–8. <https://doi.org/10.1016/j.intell.2017.04.005>
- Saqr, M., Fors, U., & Nouri, J. (2018). Using social network analysis to understand online problem-based learning and predict performance. *PloS one*, *13*(9), e0203590. <https://doi.org/10.1371/journal.pone.0203590>
- Sarathy, V. (2018). Real world problem-solving. *Frontiers in Human Neuroscience*, *12*, <https://doi.org/10.3389/fnhum.2018.00261>
- 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. DOI:10.1016/j.lindif.2012.12.011
- Schwichow, M., Croker, S., Zimmerman, C., Höffler, T., & Härtig, H. (2016). Teaching the control-of-variables strategy: A meta-analysis. *Developmental Review*, *39*, 37–63. <https://doi.org/10.1016/j.dr.2015.12.001>
- Shute, V., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, *63*, 106-117.
- Sonnleitner, P., Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R., ... Latour, T. (2012). The Genetics Lab: Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving. *Psychological Test and Assessment Modeling*, *54*(1), 54–72.

- Sonnleitner, P., Keller, U., Martin, R., & Brunner, M. (2013). Students' complex problem solving abilities: Their structure and relations to reasoning ability and educational success. *Intelligence*, *41*(5), 289-305. <https://doi.org/10.1016/j.intell.2013.05.002>
- Stadler, M., Fischer, F., & Greiff, S. (2019). Taking a Closer Look: An Exploratory Analysis of Successful and Unsuccessful Strategy Use in Complex Problems. *Frontiers in Psychology*, *10*, 1–10. <https://doi.org/10.3389/fpsyg.2019.0077>
- Stadler, M., Niepel, C., & Greiff, S. (2016). Easily too difficult: Estimating item difficulty in computer simulated microworlds. *Computers in Human Behavior*, *65*, 100–106. <https://doi.org/10.1016/j.chb.2016.08.025>
- Stadler, M., Niepel, C., & Greiff, S. (2019). Differentiating between static and complex problems: A theoretical framework and its empirical validation. *Intelligence*, *72*, 1-12. <https://doi.org/10.1016/j.intell.2018.11.003>
- Staggers, N., & Norcio, A. F. (1993). Mental models: concepts for human-computer interaction research. *International Journal of Man-Machine Studies*, *38*(4), 587-605. <https://doi.org/10.1006/imms.1993.1028>
- Taub, M., Azevedo, R., Bradbury, A. E., & Mudrick, N. V. (2020). Self-Regulation and Reflection during Game-Based Learning. In *Handbook of Game-Based Learning* (pp. 239-262). Cambridge, MS, US: MIT Press
- Trilling, B., & Fadel, C. (2009). *21st Century Skills, Learning for Life in Our Times*. San Francisco, CA: John Wiley & Sons.
- Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, *51*, 1-10.

- Urry, J. (2008). Climate change, travel and complex futures. *The British Journal of Sociology*, 59(2), 261-279.
- Vainikainen, M. P. (2014). *Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity*. Helsinki: Picaset: University of Helsinki. Department of Teacher Education Research Reports, 360.
- Wirth, J., & Funke, J. (2005). Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme. In *Problemlösekompetenz von Schülerinnen und Schülern* (pp. 55-72). VS Verlag für Sozialwissenschaften.
- Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, 10(3), 329-345.
<https://doi.org/10.1080/0969594032000148172>
- Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, 40(1), 1–14. <https://doi.org/10.1016/j.intell.2011.11.003>
- 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, 127–146.
<https://doi.org/10.1007/s10758-014-9222-8>