JDDM
Journal of Dynamic Decision Making

Open Access
Original Research

# A study on the training of complex problem solving competence

**André Kretzschmar[1,2] and Heinz-Martin Süß[2]**
[1]Institute of Psychology I, Otto von Guericke University Magdeburg, Germany and [2]ECCS research unit, University of Luxembourg, Luxembourg

**This study examined whether experience with different computer-based complex problem situations would improve complex problem solving (CPS) competence in an unknown problem situation. We had $N = 110$ university students take part in a control group study. They were trained in five different complex problem situations for up to 7 hr, and their performance was tested in a sixth complex problem situation. The data analyses revealed that the training influenced the CPS process of knowledge acquisition. However, the CPS process of knowledge application was not impacted by experience with other problem situations. Implications for the concept of CPS as a trainable competence as well as the training of CPS in general are discussed.**

**Keywords:** complex problem solving, cognitive training, transfer, flexibility training, experience, FSYS

Complex problem solving (CPS)[1] was introduced into European psychology by Dörner and colleagues (e.g., Dörner, Kreuzig, Reither, & Stäudel, 1983) and immediately attracted attention as a new cognitive ability that was applicable to real-life demands (Dörner, 1986). The handling of these real-life demands (e.g., the interconnectedness of problem areas or the dynamic development of a problem situation) has been an integral part of CPS performance and was thus included in Buchner's definition of CPS:

> *The successful interaction with task environments that are dynamic (i.e., change as a function of user's intervention and/or as a function of time) and in which some, if not all, of the environment's regularities can only be revealed by successful exploration and integration of the information gained in that process.*
> (Frensch & Funke, 1995, p. 14)

Decades later, in a time of rapid technological and scientific advances, CPS became a rising fundamental issue in science, industry, and education (e.g., Funke, 1999; Neubert, Mainert, Kretzschmar, & Greiff, 2015; OECD, 2014). For example, in the educational context, CPS has shown its utility in several respects (e.g., Greiff et al., 2013; Kretzschmar, Neubert, & Greiff, 2014; Scherer & Tiemann, 2012; Sonnleitner, Keller, Martin, & Brunner, 2013), and it is currently an important part of national and international large-scale assessments such as the Programme for International Student Assessment (PISA; OECD, 2014). Especially in this context, CPS is considered to be a cross-curricular and knowledge-based competence (OECD, 2014).

According to the understanding of CPS as a competence, CPS consequently differs in one considerable feature from related cognitive abilities such as intelligence (see Süß, 1996; Wittmann & Hattrup, 2004; Wüstenberg, Greiff, & Funke, 2012). Whereas cognitive abilities are relatively stable over time and are not viewed as trainable, competencies are per definition modifiable through interventions (Weinert, 2001). This article adopts the perspective of viewing CPS as a competence and aims to examine its trainability.

## The Training of Complex Problem Solving Competence

In general, the training and transfer of different cognitive achievements has been an exciting research topic that is highly relevant to real life. For example, recent studies have demonstrated that training with video games may improve cognitive performance outside the game context (e.g., basic visual attention or executive control; e.g., Anguera et al., 2013; Green & Bavelier, 2003; Strobach, Frensch, & Schubert, 2012). Although the findings and especially the transfer to untrained tasks have been critically discussed (e.g., Boot, Blakely, & Simons, 2011), the rationale behind such cognitive training approaches is that training does not improve performance on only a single task, but rather that training improves achievements on other cognitive tasks or with regard to real-life criteria, too.

Naturally, a CPS competence training should also reflect such an approach. This means that the training should not be limited to better performance only in the problem situations that are trained but should rather manifest in better performance in unknown problem situations. Surprisingly, after almost 40 years of research on CPS, the inclusion of CPS as a competence in educational large-scale assessments (OECD, 2014), as well as a steadily increasing interest in the trainability of CPS competence (Dörner, 1976; Funke, 2003), there is a remarkable lack of research that has provided an understanding and empirical examinations of CPS competence training — especially with regard to the transfer of CPS skills to unknown problem situations. This is all the more astonishing as more or less explicit recommendations for how to increase an individual's CPS competence and how to change school practices and educational polices in order to foster CPS competence have been documented (see OECD, 2014). Therefore, the purpose of this study was to reduce this gap in research and to shed some light on the issue of the extent to which CPS competence is trainable.

**Corresponding author:** André Kretzschmar, ECCS unit, University of Luxembourg, 11, Porte des Sciences, 4366 Esch-Belval, Luxembourg; Phone: +352-466644-9245; Fax: +352-466644-5376.
e-mail: kretzsch.andre@gmail.com; ORCID ID: 0000-0002-7290-1145.

## The Flexibility Training Approach

Deliberate practice is defined as engagement in activities that are specifically designed to improve performance in a domain (Meinz & Hambrick, 2010). In this sense, Dörner (1989) proposed the flexibility training approach in accordance with the assumption that a CPS competence training should also have an effect on untrained complex problems. The aim of the flexibility training is to develop general problem solving knowledge (GPSK) about how to solve complex problems. GPSK (or heuristic knowledge; Schaub & Strohschneider, 1992) is knowledge about the need to explore the problem (e.g., to acquire situation-specific knowledge), how to conduct interventions (e.g., careful interventions in unstable systems), and how to reach goals (e.g., how to effectively organize a series of interventions). GPSK can therefore be considered meta-problem-solving knowledge that is applicable across different situations. According to Dörner and other reseraches, GPSK can be developed by gaining experience with different problem situations (Dörner, 1976; Schaub & Strohschneider, 1992; Strohschneider, 1990). This means that comprehensive experience with problem situations involving heterogeneous demands leads to the successive abstraction of problem solving procedures. Due to this abstraction process, it can be assumed that problem solving knowledge will be less dependent on concrete problem situations and, thus, generally applicable (Weinert & Waldmann, 1988). For example, imagine that you have just bought a computer with a new operating system (e.g., Ubuntu Linux). A problem situation may develop if you want to install a software program. Your (situation-specific) knowledge about how to install new software in your old operating system (e.g., Microsoft Windows) would be less useful because the procedures for installing software differ between the two operating systems. Therefore, an appropriate way to proceed would be first to acquire knowledge about the operating system's installation procedure by consulting the help pages. Consulting the help pages or, more generally, knowing how to acquire essential information about a problem is an example of GPSK. Experience with different problem situations would lead to the GPSK: "If I do not know the specific procedure for solving the problem, then I should use the support that is offered to get the information." Consequently, if your office computer has another operating system (e.g., OS X), neither your specific knowledge about Linux nor about Microsoft Windows will be sufficiently helpful. However, the GPSK of *using the help pages* will increase your chances of successfully installing new software on your office computer too. Moreover, having experience with different operating systems would also increase a person's knowledge about common solutions (i.e., how to apply knowledge about software installation). Although the specific procedures differ between the operating systems, the problem solver might realize the similarity of the principal steps (e.g., start the procedure, configure some features, check the success of the procedure). Experience with different problem situations should, therefore, increase GPSK in numerous ways. However, such experience might not be limited to the contexts of software installation in different operating systems but should also be useful in different contexts (e.g., using an unfamiliar mobile phone).

Basically, Dörner's (1989) flexibility training approach follows the assumption that a person's experience with different complex problem situations will help the person develop a higher GPSK, which leads to a better CPS performance in an unknown problem. However, as always in training contexts, the essential question here is how to teach GPSK, especially with regard to its transferability to unknown problem situations. Previous research has shown that the direct teaching of GPSK (e.g., general problem solving strategies) is rather unrewarding, whereas a learning-by-doing approach (Anzai & Simon, 1979) seems to be generally more efficient (e.g., Friedrich & Mandl, 1992; Putz-Osterloh, 1988; Stern, 1993). Consequently, the flexibility training should be primarily based on direct interactions between the problem solver and different complex problem situations (i.e., learning-by-doing).

In summary, the flexibility training approach followed the principle of developing GPSK by gaining experience with different problem situations. Whereas specific problem solving knowledge (e.g., installing a software program in a new operating system) is less advantageous in unknown situations, the use of GPSK (e.g., how to acquire knowledge and follow the common procedures) promotes the solving of new and unknown problem situations.

## Previous Empirical Findings on CPS Competence Training

The impact of experience on complex problem situations has been examined with expert-novice comparisons (e.g., Putz-Osterloh, 1987; Schaub & Strohschneider, 1992). However, to our knowledge, no studies have aimed to investigate the training of general CPS competence (i.e., the development of GPSK). In fact, the vast majority of studies that have had the goal of increasing CPS competence have used only a single complex problem situation without considering whether competence in CPS could be transferred to novel problem situations. The consistent, albeit quite trivial, finding of these studies is that people who are trained in a specific problem situation perform better when confronted with the same one (Funke, 2006). Unfortunately, such studies cannot contribute to answering the question of whether CPS competence is trainable. The effects of practice have been shown several times in different contexts (e.g., Kulik, Kulik, & Bangert, 1984), but this does not necessarily imply an improved competence.

With regard to the few studies that have effectively focused on effects that can be successfully transferred to new problem situations, such studies have used only two problem situations (i.e., one for the training and one for evaluating the transfer) with limited training time (Bakken, 1993; Jensen, 2005; Putz-Osterloh & Lemme, 1987). For example, in Jensen's (2005) study, people were trained with the rabbit-and-fox task, and the transfer performance was evaluated with the reindeer-and-lichen task. The results indicated that people were able to learn from experience with complex problem situations and use that knowledge in a new problem situation (Jensen, 2005). However, the two tasks were quite similar, and thus, the results could not be used to determine whether the participants had developed GPSK that could be applicable to a less similar problem situation or whether they had acquired only task-specific knowledge instead.

Some indirect evidence for the development of GPSK through experience has come from research on CPS assess-

[1]There are several synonyms for CPS in the literature, for example, dynamic decision making (e.g., Gonzalez, Thomas, & Vanyukov, 2005), interactive problem solving (e.g., Fischer et al., 2015), dynamic problem solving (e.g., Greiff, Wüstenberg, & Funke, 2012), and creative problem solving (e.g., OECD, 2014).

ment tools using the multiple-item approach (e.g., MicroDYN; Greiff et al., 2012). These tools evaluate the use of exploratory behavior in a sequence of more or less different problem situations. With these tools, problem solvers are given no feedback concerning their exploration behavior and learn only from their experience with the tasks. The empirical findings (e.g., Schweizer, Wüstenberg, & Greiff, 2013; Wüstenberg et al., 2012) have clearly shown an increase in the use of the specific exploration strategy VOTAT (i.e., vary one thing at a time; Vollmeyer, Burns, & Holyoak, 1996) across the task sequences. Although the use of VOTAT was not taught during the assessment and no feedback was provided, problem solvers discovered the advantages of that strategy and used it more often at the end than at the beginning of the assessment. However, the tasks presented in MicroDYN are again highly similar, and thus, it is unknown whether problem solvers would be able to apply the knowledge they learned (i.e., use of the VOTAT strategy) to solve other, less familiar CPS tasks.

To sum up, the previous findings point to the possibility of learning through experience with problem situations as well as to the ability to transfer one's experience to similar complex problem situations. However, the question of whether experience with different problem situations leads to an improved CPS competence that is applicable to new and unknown problem situations is still open.

### The Present Study

The aim of the current study was to examine the extent to which CPS competence is trainable. Therefore, we chose Dörner's (1989) flexibility training approach, which can theoretically be applied to develop GPSK. We specifically hypothesized that problem solvers who were allowed to gain experience from different problem situations would perform better in an unknown complex problem situation than a control group.

## Method

### Participants

One hundred fifty-nine students from a German university participated in the experimental study. Participants in both the training and control groups were given the same incentives. In detail, psychology students were given partial credit for course requirements, whereas all other participants took part in a book raffle. Furthermore, all were given individual feedback on their performance. One hundred ten students completed the entire training and were included in the analyses. A screening of the available information indicated selective dropout. In detail, nonparticipants were mainly nonpsychology students (92%). This might indicate that the incentive of partial credit for course requirements was stronger than the book raffle incentive. Further evidence of selective dropout came from significantly better performances on some cognitive measures by the participants in comparison with the nonparticipants. The possible consequences of this selection process are discussed below. Of the final sample, 47% studied psychology, 22% mechanical engineering, 17% economics, and the rest another field of study. The mean age of the final sample was 23.28 years ($SD = 4.01$), and 49% were female. Gender was equally distributed (50% female) within each group.

### Design and General Procedure

Participants were equally recruited, meaning that the study was described as having a study length of up to 12 hr for all participants. Half of the participants completed the training, whereas the others provided a no-contact control group. Participants were randomly allocated to the training or control group when they registered for the study. With regard to group equivalence, we screened for important determinants of CPS competence (e.g., Bühner, Kröner, & Ziegler, 2008; Greiff, Kretzschmar, Müller, Spinath, & Martin, 2014; Süß, 1996; Wittmann & Hattrup, 2004). Therefore, we used several subtests from a comprehensive test of the Berlin Intelligence Structure Model (BIS test; Jäger, Süß, & Beauducel, 1997; for a description in English, see Süß & Beauducel, 2015) to measure processing capacity (i.e., reasoning, 9 tasks) and perceptual speed (9 tasks). Three different tasks from Oberauer, Süß, Wilhelm, & Wittmann (2002) and Sander (2005) were used to ascertain working memory capacity. In addition, we assessed computer skills with the short version of the computer knowledge questionnaire START-C (Wagener, 2007), a questionnaire on computer experience, and a computer-based simple reaction-time task by Sander (2005). Finally, we administered a new questionnaire (18 multiple-choice items) to gather prior domain-specific knowledge of the complex problem situation that was used to evaluate the success of the training.

Each group was tested separately. In the first session, participants filled out a questionnaire asking for personal data and were tested for baseline measures in groups of 10 to 20 persons. In the second session, the control group (on the following day) and the training group (after 1 week) were tested for CPS and domain-specific knowledge. All participants spent about 2.5 hr in each test session.

### Training Intervention Phase

**Material.** On the basis of the flexibility training approach, we used five different microworlds (i.e., computer-based complex problem solving situations) as training tools in order to provide problem-situation demands that were as heterogeneous as possible. The selection of microworlds was guided by a literature review on training purposes, but no computer program that was specifically designed for training was available. Therefore, we consulted the CPS research literature to choose the following microworlds, which differed sufficiently among themselves (as a prerequisite for the flexibility training) and with regard to the preselected microworld for the training evaluation (i.e., no equal semantic embedding, no equal user interface, etc.). In general, there is not yet a convincing way to objectively compare microworlds. Wagener (2001) provided a very elaborate and comprehensive framework with 43 features to classify a broad range of microworlds. Unfortunately, most of the features are not applicable to each microworld and, even more important, the classification is rather subjective. However, Table 1 provides an overview of the formal features that are roughly based on Wagener's framework in order to provide additional information about the microworlds we chose and their comparability.

**ColorSim.** The microworld ColorSim (Kluge, 2008; see Figure 1) functioned as the initial simulation and was intended to get participants started. It is easy to understand and, thus, it provided a good demonstration of the basic principles of solving complex problem situations. ColorSim has no real-world embedding. Problem solvers have to manipulate three slide switches to reach the target values of three output parameters. The different levels of the microworld are implemented through an increasingly complex network comprised of the different variables. We used the easiest level as well as a moderate level according to Kluge (2008). In our study, each level consisted of nine tasks. ColorSim is turn-based and does not have a strict time limit.

**K4.** The aim of K4 (Wagener, 2001; see Figure 2) is to manage a publishing house, in particular, to produce and sell magazines. Thus, the problem solver has to control the price, the quality, the circulation, and so forth depending on demand and customer satisfaction. This microworld has three levels realized through the numbers of variables and manipulable parameters and the interconnections between them. We used all three levels; a level took between 15 and 75 min.

**PowerPlant.** The microworld PowerPlant (Wallach, 1998; see Figure 3) provides a realistic simulation of a coal-fired power station. Problem solvers have to control the system by manipulating two actuating elements: the supply of coal and the opening of a steam valve. They have to follow a target demand energy curve. The difficulty between the sessions changes with different target profiles of the demand curve. In our study, each session consisted of an introductory phase (15 decisions) followed by a performance phase of approximately 20 min.

**Tailorshop.** We used the version of this microworld presented by Süß (1996; see Figure 4). In the Tailorshop, problem solvers have to obtain economic success by manipulating prices, buying raw material, controlling wages, and so forth. Two different Tailorshop start conditions were used. In the current study, each of the two sessions consisted of 2 (simulated) months of tutorials, an exploration phase of 30 min, and a turn-based performance phase of 12 months.

**Networked Fire Chief.** In Networked Fire Chief (Omodei & Wearing, 1998; see Figure 5), problem solvers have to fight fires and coordinate their task forces. The task is to move units to the location of the fire while simultaneously managing water resources. For this study, we developed a new level that considered the special demands of CPS (e.g., assigning priorities, observing different targets). The duration of the level was approximately 20 min.

**Procedure.** At the end of the first session, people in the training group received a short introduction to the technical details of the computer training and completed the training individually at home. Table 2 shows the training plan, consisting of 10 sessions with five different microworlds. Each microworld had to be handled at least twice with each including an exploration phase and a performance phase. In the exploration phase, the problem solvers had no task-related objectives. Thus, subjects had the opportunity to acquire task-related knowledge, check out different strategies, learn from their mistakes, and evaluate the process without the pressure to perform (see Osman, 2010; Vollmeyer & Funke, 1999). No additional hints or feedback were given during the training phase; that is, the training was exclusively based on learning-by-doing (see Anzai & Simon, 1979). Furthermore, the training was designed with increasing difficulty and complexity and with an alternating presentation of microworlds (see Goettl, Yadrick, Connolly-Gomez, Regian, & Shebilske, 1996).

The training plan was designed with a time period of 1 week and a daily training between 40 and 90 min. On average, participants spent about 6.5 hr on the training. However, this is only a rough estimate of the total training time because the training software (see below) did not provide precise time-on-task information.

We designed a training system that permitted the participants to complete all of the training sessions at home via the Internet at any time. To do this, every trainee received a training manual with general and software-related information describing the basic functions used in each microworld. For every training session, the participants had to log into an online environment with their personal log-in data[2]. Next, a software program that coordinated and recorded the entire training automatically started a microworld according to the individual's training progress. No training session could be skipped or repeated, but under certain circumstances (e.g., a disrupted Internet connection), it was possible to start only the previous session again. After completing the last session, no further training was possible.

[2]The microworlds used for the present training were initially developed to run on a local computer system. In order to provide training that could be applied online, we emulated such a local computer system with the help of virtual machines. Participants used the Remote Desktop Protocol (RDP) to connect to the virtual machines via the Internet.

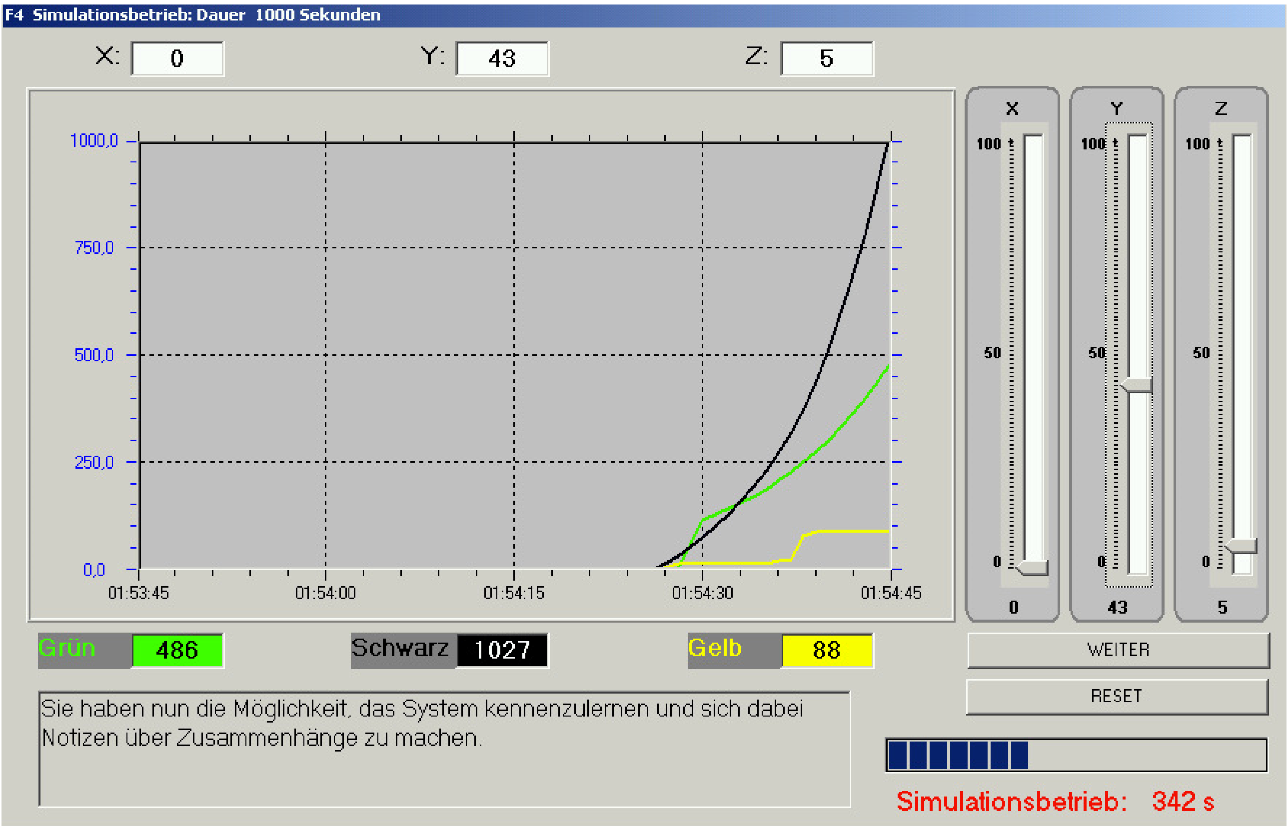


**Figure 1.** Screenshot of the ColorSim microworld.

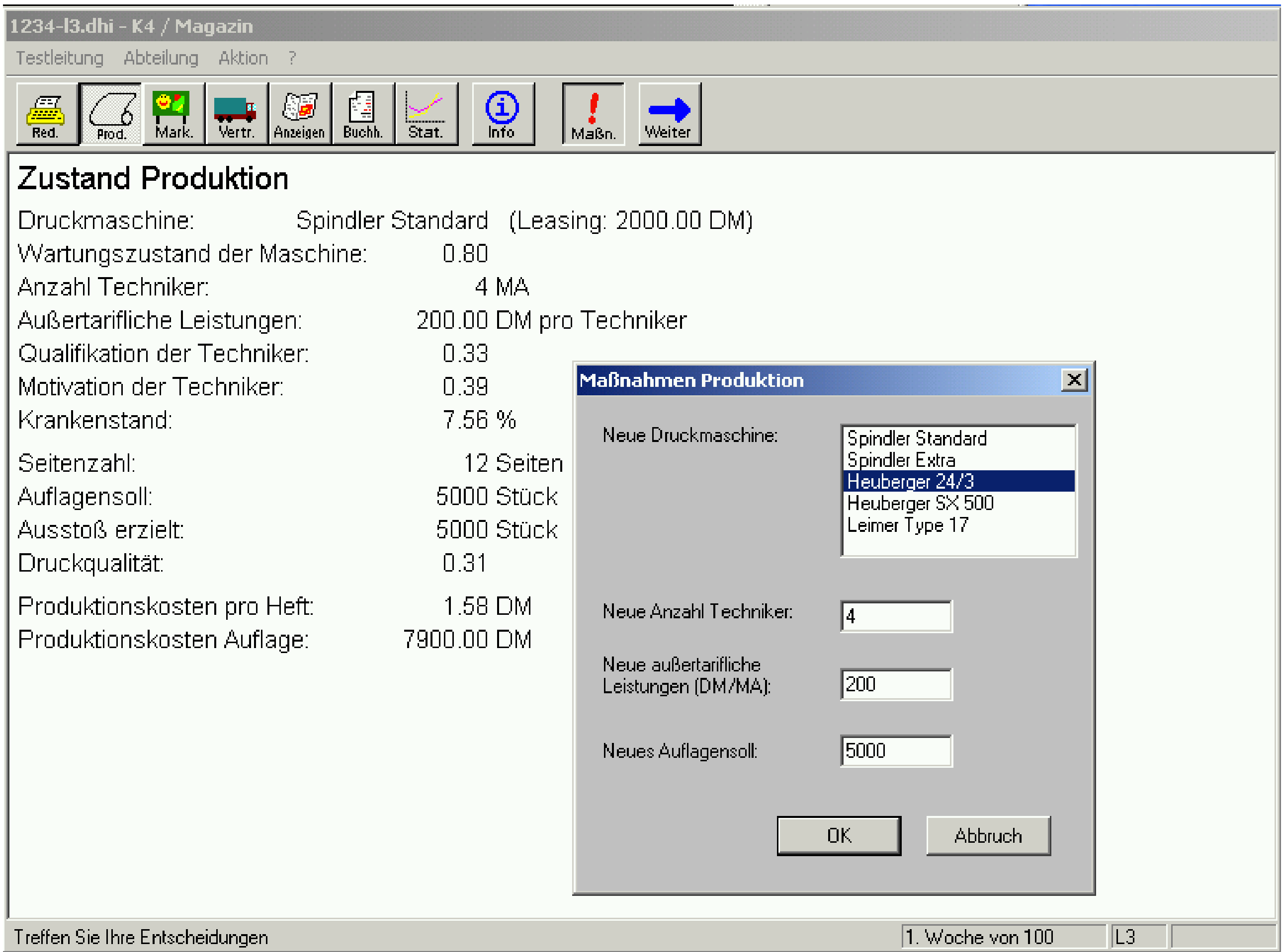


**Figure 2.** Screenshot of the K4 microworld.

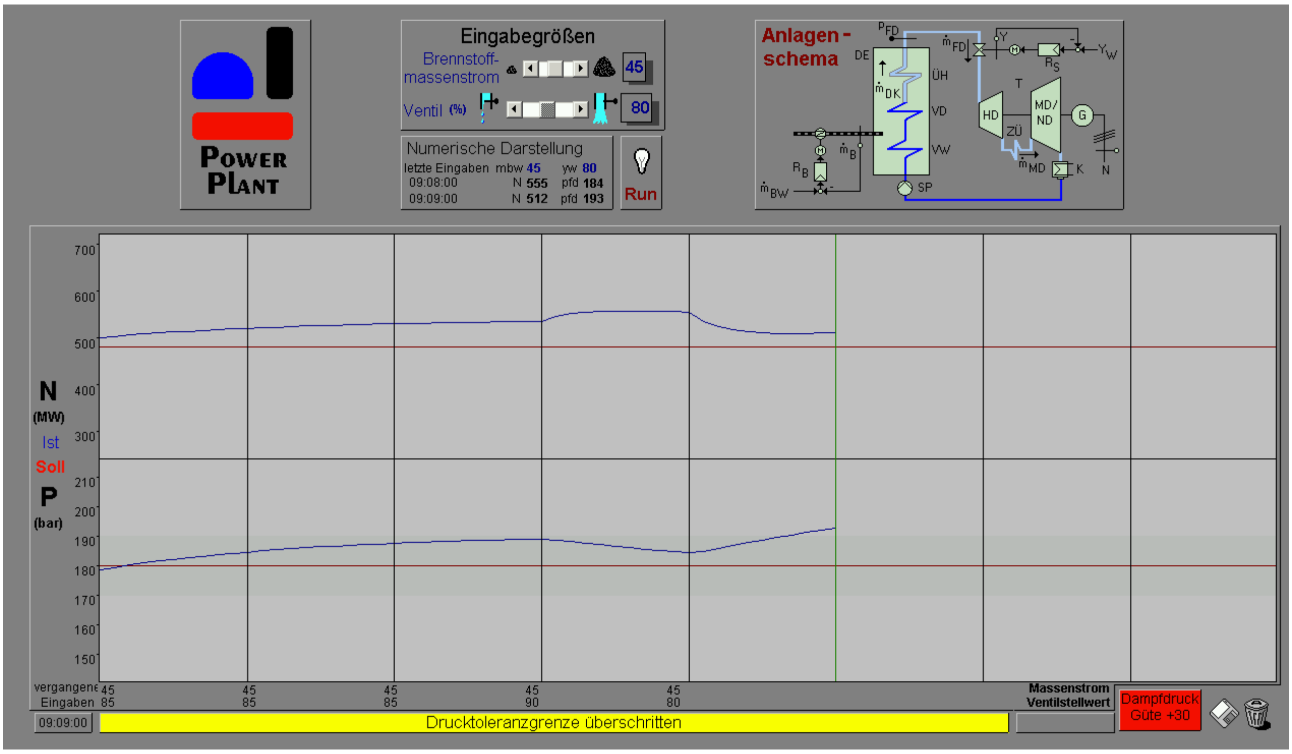


**Figure 3.** Screenshot of the PowerPlant microworld.

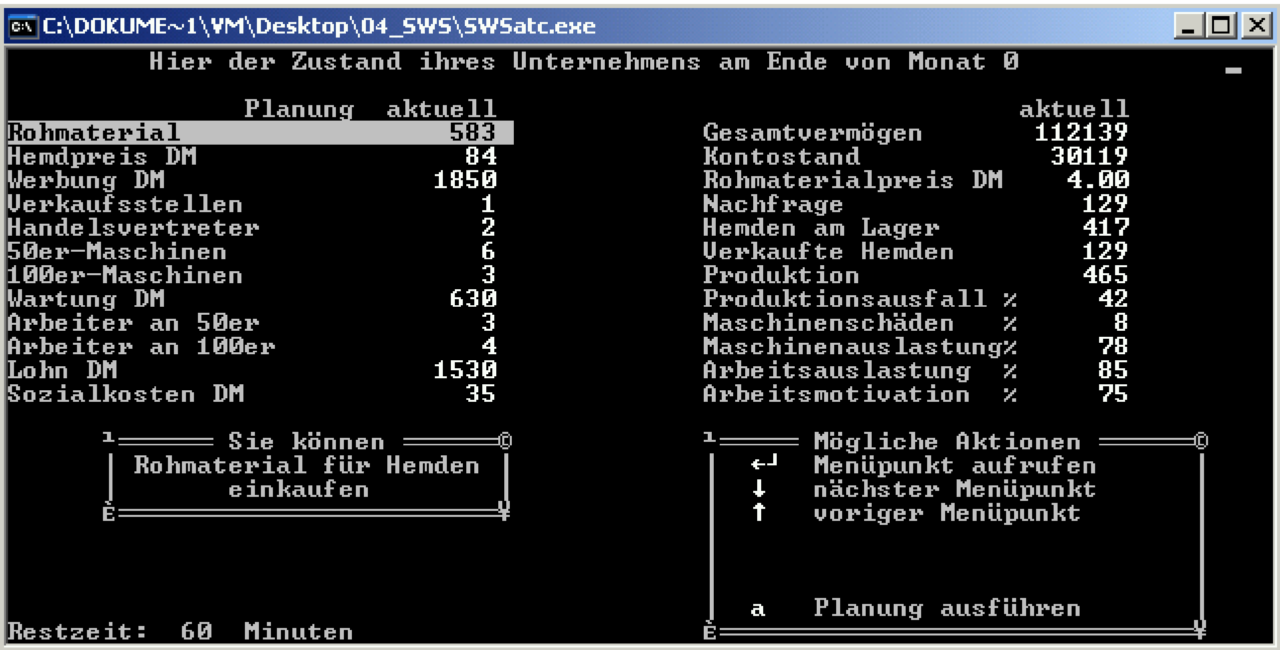


**Figure 4.** Screenshot of the Tailorshop microworld.

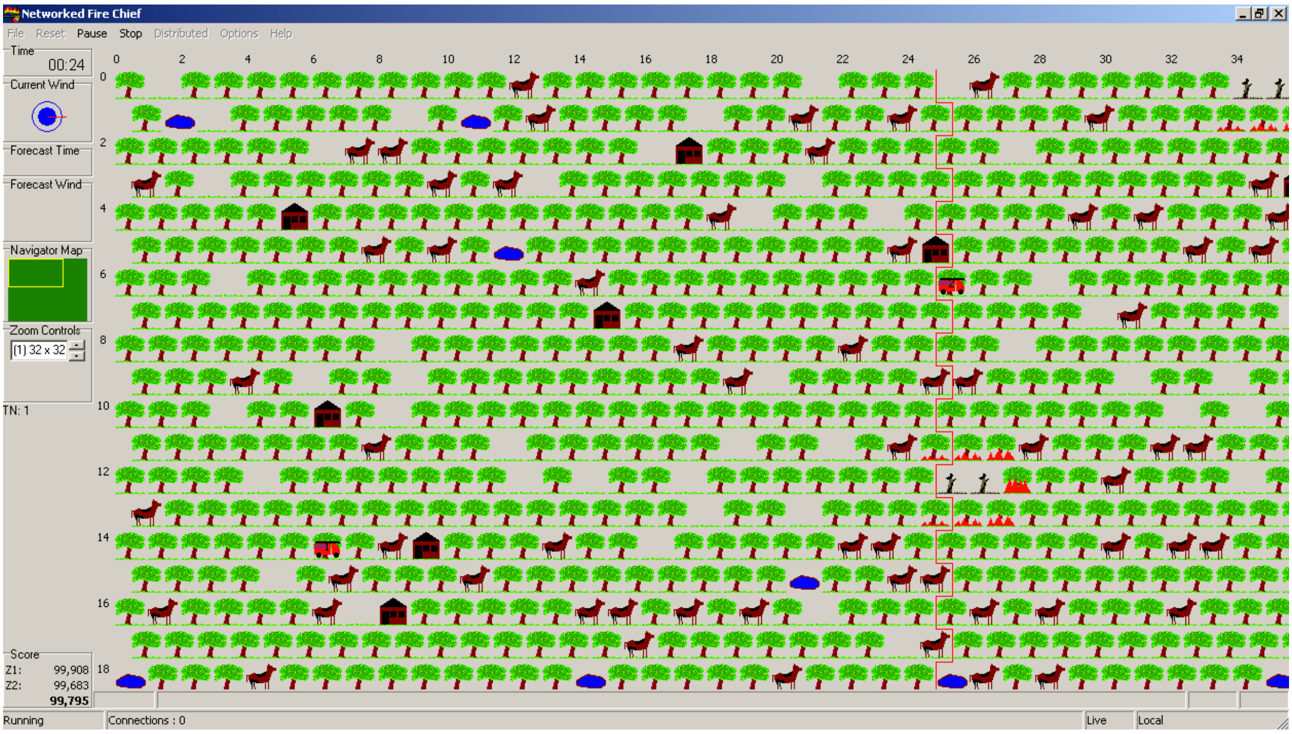


**Figure 5.** Screenshot of the Network Fire Chief microworld.

**Table 1.** Selection of formal features of microworlds roughly based on Wagener's (2001) framework.

| Features | ColorSim | K4 | PowerPlant | Tailorshop | Networked Fire Chief | FSYS |
|---|---|---|---|---|---|---|
| Semantic embedding | No/abstract | Publishing company/ management | Energy production/ engineering | Clothing factory/ management | Fire fighting | Forestry/ management |
| Impact of prior knowledge | No | Moderate | Low-moderate | Moderate | Low | Low |
| Content presentation | Numerical, in part figural | Numerical | Figural, in part numerical | Numerical | Figural | Numerical, in part figural |
| Turn-based | Yes | Yes | No | Yes | No | Yes |
| Time limit | No | No | Yes | Yes | Yes | No |
| Real-time simulation | No | No | No | No | Yes | No |
| Number of variables | 6 | 23/31/56 | 11 | 24 | - | 85 |
| Connections between variables | Linear, multiplicative, logistic | Linear, multiplicative, logistic | Differential equations | Linear, exponential | - | Linear, exponential, logistic |
| Eigendynamics | Yes | Yes | No | Yes, exponential | Yes | Yes |
| Time delay of feedbacks | Yes | Yes | Yes | No, exponential | No | Yes |
| Hidden/indirect effects | Yes | Yes | Yes | Yes, exponential | Yes | Yes |
| Random influences | No | No | No | Yes (pseudo) | No | No |

*Note.* Impact of prior knowledge was estimated. All other information is based on the literature and test descriptions.

### Training Evaluation Phase

**Material.** We used the microworld FSYS 2.0 (Wagener, 2001; see Figure 6) to measure the training outcome. FSYS is a reliable and well-validated CPS competence assessment tool that is based on Dörner et al.'s (Dörner, 1986; Dörner et al., 1983) theoretical complex problem solving framework. The semantic embedding of FSYS is composed of a forest enterprise. However, to achieve good performance, no previous knowledge is required because the program uses fantasy names, an integrated information system, and very general cause-effect relations. The aim was to achieve economic success within 50 simulated months. Thereby, the problem solver had to manage five different wooded areas by planting and lumbering trees, fertilizing, and fighting vermin. Wagener and Wittmann (2002) demonstrated that FSYS offered incremental predictive validity beyond general intelligence with regard to job-related performance indicators (e.g., in-basket exercises, case studies). Stadler, Becker, Greiff, and Spinath (2015) reported that FSYS explained incremental variability in university grade point average when they controlled for high-school GPA and a short test of general intelligence, even though their sample size was rather small.

Furthermore, we used Wagener's (2001) questionnaire to assess FSYS-specific knowledge after the participants completed the 50 simulated months. The questionnaire addressed all relevant fields in the microworld and was composed of 11 multiple-choice questions of factual and action knowledge. For each question, five different answer options were provided (i.e., four distractors in addition to the right answer). Example questions are: "Which tree has the highest yield?" and "A forest is infested by vermin XY. Which procedure would you apply?" Wagener (2001) reported an internal consistency of .41, but this is an inaccurate estimate of the reliability because of the heterogeneity of the scale.

**Dependent variables.** The total score on the FSYS-specific knowledge test (Wagener, 2001) was used as an indicator of the process of knowledge application. Each item was scored dichotomously, resulting in a sum score that ranged from 0 to 11. For the process of knowledge application, we used the control performance in FSYS[3], that is, the total amount of property amassed by the end of the simulation (original name SKAPKOR). According to the standard scoring procedure, the property value was logistically transformed into a scale ranging from 0 to 100 such that higher values indicated better performance (Wagener, 2001).

Wagener (2001) reported substantial manifest correlations ($r$ = .41 to .44) between knowledge acquisition (i.e., knowledge test) and knowledge application (i.e., control performance) for FSYS. The correlation in the present study was comparable ($r$ = .53, $p < .001$) and fell within the range commonly obtained for these two processes (Goode & Beckmann, 2010). The CPS processes of knowledge acquisition and knowledge application were analyzed separately to examine potential differential effects of the training.

## Results

An alpha level of .05 was used for all statistical tests. All significance tests were one-tailed except when noted otherwise. In addition to the significance levels, Cohen's (1988) effect sizes are reported. The sample size was adequate for detecting at least medium-sized effects ($d = 0.5$) with a power of .80 for simple mean comparisons (Faul, Erdfelder, Lang, & Buchner, 2007).

Participants' gender (50% female within each group), age, perceptual speed, working memory capacity, computer skills, and domain-specific knowledge did not differ between the two groups ($t_{\text{s}} \leq 1.5$, $p_{\text{s, two-tailed}} > .10$). However, there was a significant difference in reasoning, $t(108) = 2.65$, $p_{\text{two-tailed}} = .01$, $d = 0.51$, with superior performance in the training group. Therefore, group equivalence was not completely achieved, and all further analyses were additionally run with reasoning as a covariate. The descriptive statistics are shown in Table 3.

We had expected better CPS performance in the training group than in the control group. We first analyzed CPS competence with regard to the process of knowledge acquisition. The training group ($M$ = 5.91, $SD$ = 1.93) showed better performance than the control group ($M$ = 4.89, $SD$ = 1.79). The difference was statistically significant, $t(108) = 2.87$, $p_{\text{one-tailed}} < .01$, with a moderate effect size ($d = 0.55$). The results did not change even when reasoning was controlled for.

In a second step, we analyzed CPS competence with regard to the process of knowledge application. The two groups showed almost identical performance in controlling FSYS. The mean score for the training group was 58.36 ($SD$ = 20.77) and was 57.14 ($SD$ = 24.86) for the control group. Thus, there was no statistically significant difference, $t(108) = 0.28$, $p_{\text{one-tailed}} = .39$, $d = 0.05$. Again, controlling for reasoning did not change the findings.

In summary, our findings only partly supported our hypothesis. After the flexibility training, the training group participants were significantly better at acquiring knowledge in comparison with the control group participants. However, no significant difference was found in knowledge application (i.e., in solving the problem).

## Discussion

Solving nonroutine problems is highly relevant in our rapidly changing world, and consequently, CPS competence plays an important role, especially in the educational context (OECD, 2014). Although CPS competence is considered to be trainable (OECD, 2014),

[3] FSYS also provides additional behavioral scales that were not used in this study. Most of the behavioral scales did not exhibit acceptable psychometric quality, and thus, we could not interpret them without reservations (see Wagener & Conrad, 2001).


**Figure 6.** Screenshot of the Tailorshop microworld.

**Table 2.** Training program.

| Session | Microworld | Version/Level of difficulty | Estimated duration |
|---|---|---|---|
| 1 | ColorSim | 1 | 35 |
| 2 | ColorSim | 2 | 45 |
| 3 | K4 | 1 | 35 |
| 4 | PowerPlant | A | 30 |
| 5 | Tailorshop | A | 90 |
| 6 | K4 | 2 | 60 |
| 7 | PowerPlant | B | 20 |
| 8 | Fire Chief | — | 20 |
| 9 | K4 | 3 | 75 |
| 10 | Tailorshop | B | 60 |

*Note.* Each session was composed of an exploration phase that did not involve a performance evaluation and a performance phase. The letters A and B indicate different versions; the numbers indicate the level of difficulty. The estimated duration in minutes is based on information from the associated literature.

**Table 3.** Descriptive statistics (means, confidence intervals, standard deviations) for both groups.

| Measure | Control group | | | Training group | | | |
|---|---|---|---|---|---|---|---|
| | M | 95% CI | SD | M | 95% CI | SD | $\omega$ |
| Age | 23.96 | [22.73, 25.19] | 4.55 | 22.80 | [21.90, 23.70] | 3.34 | — |
| Reasoning | -0.14 | [-0.30, 0.01] | 0.58 | 0.13 | [0.00, 0.27] | 0.51 | .88 |
| Perceptual speed | -0.07 | [-0.25, 0.10] | 0.65 | 0.10 | [-0.07, 0.27] | 0.63 | .79 |
| Working memory capacity | 0.04 | [-0.50, 0.57] | 1.97 | 0.58 | [0.07, 1.09] | 1.88 | .89[1] |
| Computer knowledge | 18.36 | [17.37, 19.36] | 3.67 | 18.71 | [17.73, 19.68] | 3.6 | .79 |
| Computer experience in years | 10.27 | [9.5, 11.04] | 2.84 | 10.51 | [9.54, 11.47] | 3.57 | — |
| Simple reaction-time | 255.08 | [249.92, 260.23] | 19.07 | 256.24 | [251.10, 261.38] | 19.01 | .92 |
| Domain-specific prior knowledge | 8.18 | [7.52, 8.85] | 2.46 | 8.40 | [7.56, 9.24] | 3.09 | .59 |
| CPS: Knowledge acquistion | 4.89 | [4.41, 5.38] | 1.79 | 5.91 | [5.39, 6.43] | 1.93 | .53 |
| CPS: Knowledge application | 57.14 | [50.42, 63.86] | 24.86 | 58.36 | [52.74, 63.97] | 20.77 | — |

*Note.* For reasoning (Z scores), perceptual speed (Z scores), working memory capazity (Z scores), computer knowledge and domain-specific prior knowledge, the sum scores of the single tasks were used. For simple reaction-time the mean score was used.
[1]Only for the subtask reading span; for the other two tasks (dot span and memory updating numerical) a single total score was used.
$\omega$: McDonald's Omega.

empirical support for this supposition has been rather weak. The purpose of the current study was to deepen the understanding of the extent to which CPS competence can be improved through a training intervention. The findings of this study showed that CPS competence was only partly influenced by flexibility training. Whereas training significantly improved the CPS process of acquiring relevant knowledge about the problem situation, the process of knowledge application was not affected. Aside from the differential effects of training on the two CPS processes, the general implications for the trainability of CPS competence will be discussed.

According to Fischer, Greiff, & Funke (2012), CPS involves at least two consecutive processes. First, the problem solver has to acquire problem-specific knowledge (e.g., how to install software in a new operating system), and second, the problem solver has to apply this knowledge in order to solve the problem (e.g., complete the software installation process). Apparently, handling different problem situations improves the first step involved in solving complex problem situations. That is, in our study, trained problem solvers were able to acquire more knowledge about an unknown problem situation. According to the flexibility training approach, this finding can be interpreted as improvements in GPSK concerning the utility of knowledge acquisition. When frequently confronted with different problem situations where no prior knowledge is applicable, the problem solver recognizes that he or she must acquire knowledge about a problem in order to solve it. Moreover, processes through which people can effectively acquire knowledge (e.g., different exploration strategies such as VOTAT; Vollmeyer et al., 1996) might also be improved through experience with different problem situations. This issue is especially important when considering problem solving in real life. Real-life problems differ widely and, thus, situation-specific knowledge might be not available for each problem (e.g., an unfamiliar mobile phone or a new operating system). However, when problem solvers can rely on GPSK, their chances of acquiring situation-specific knowledge that will help them solve the problem (e.g., through comprehensive exploration) increase. In this sense, flexibility training can be considered successful with regard to the prerequisite of actually solving complex problems, that is, acquiring knowledge about an unknown problem situation.

On the other hand, GPSK consisting of how to apply the acquired information in order to solve the problem does not seem to be improved by the training. This finding is crucial when evaluating CPS training in general. Although acquiring relevant information is necessary for solving an unknown problem, the actual goal is to solve the problem. Returning to the introductory example, knowing how to obtain information that explains how to install a software program but to still be unable to actually install it will not lead to a satisfactory outcome. The reason for the lack of impact from the CPS training might lie in different issues.

In general, the flexibility training is aimed at improving CPS competence independent of a particular problem situation in terms of GPSK. However, it might be the case that the process of knowledge application is rather situation-specific. This means that every problem situation primarily requires specific knowledge about how to act in the situation, and thus, GPSK that is independent from the actual problem situation might play only a minor role in a person's ability to actually solve the problem. The importance of previous knowledge in a specific problem situation (see Süß, 1996) might support this perspective. Furthermore, our flexibility training was less focused on addressing the special demands of knowledge application (e.g., careful interventions in unstable systems). In fact, trainees were encouraged to extensively explore each problem situation during the training to understand how to solve the problem. Although the process of knowledge application (i.e., actually solving the problem) was covered in each training session, it is unknown whether trainees primarily used the performance phase to solve the problem (e.g., to be a successful Tailorshop manager) or whether they used the performance phase to also learn more about the problem situation. The latter would mainly lead to an increase in GPSK related to the process of knowledge acquisition rather than the process of knowledge application. In this respect, it is important to note that the training was an unguided training; that is, we provided no feedback, no explicit teaching of problem solving strategies, or any similar guidance. Instead, the training was completely based on the learning-by-doing approach (Anzai & Simon, 1979). A more guided training (e.g., emphasizing training goals, problem solving phases, and processes, etc.) might lead to an increase in the training effect also for the process for knowledge application. Therefore, training studies that are tailored to address the specific demands of applying knowledge combined with a more explicitly guided training approach might be able to shed some light on the question of whether GPSK can be improved with respect to the problem solver's ability to actually solve the given problem.

Another explanation for the findings can be found in assessments of the CPS processes. The assessment of knowledge application often has limitations in terms of reliability; that is, the assessment using microworlds is actually a single-item measure with limited reliability (Beckmann & Goode, 2014; Greiff et al., 2012; Süß, 1999). Although the estimated internal consistency of FSYS is rather high (.80; Wagener, 2001), its test-retest reliability and parallel-test reliability remain unknown. Thus, limitations in the reliability of FSYS might prohibit the ability to use it to detect the success of the training with respect to the process of knowledge application. Recently developed measurement tools focusing in particular on psychometric criteria (e.g., Neubert, Kretzschmar, Wüstenberg, & Greiff, 2014; Sonnleitner et al., 2012) might be used to resolve this issue as they provide reliable scores for knowledge acquisition and knowledge application.

## Limitations and Recommendations for Further Research

Some limitations of this study need to be discussed, especially with respect to future CPS training studies. First, the no-contact control group did not complete any pseudo-training. We had expected that pseudo-training would not have affected the training outcome. The few previous training studies had not shown substantial transfer effects regardless of whether participants had received training of any kind or a control intervention. Furthermore, the intensive use of computers by many students has become commonplace (Prensky, 2001), which was confirmed by the analyses of the background questions used in this study. Thus, we did not expect training to produce any improvements due solely to the use of computers. Finally, both groups were informed up front that the study duration could be up to 12 hr so that the motivation in both groups could be assumed to be equal. In summary, an additional intervention for the control group seemed dispensable, but it might be beneficial to include an active control group in future research.

Furthermore, despite the use of random assignment, group equivalence was not achieved. The training group outperformed the control group in reasoning, which has repeatedly been shown to be a substantial predictor of CPS performance (e.g., Kretzschmar, Neubert, Wüstenberg, & Greiff, 2016; Sonnleitner et al., 2013; Süß, 1996; Wittmann & Süß, 1999; Wüstenberg et al., 2012). Although we statistically controlled for the difference in reasoning on the pretest, we have to consider that the training group may have benefitted from their higher reasoning ability (Matthew effect; e.g., Walberg & Tsai, 1983). However, we can only speculate about the reasons for the differences between the groups. Although random assignment to the training or control group was applied, it does not ensure group equivalence, especially when sample sizes are small or moderate (e.g., Saint-Mont, 2015). Thus, we have to consider the possibility that even additional group differences (e.g., in motivation) may have influenced participants' performance in our study.

Another issue involves the sample characteristics. The participants were recruited from a sample with above-average cognitive performance. Although university students are often used in psychological studies, such findings should be generalized only with caution (Henrich, Heine, & Norenzayan, 2010). In fact, the selective dropout in the present study even reduced the heterogeneity of the sample with notable consequences. That is, the effect of the present training may have been underestimated due to a ceiling effect for the participants with above-average cognitive abilities. Therefore, it is possible that the effect of the training would be stronger in a less selective sample. Future CPS research should aim to avoid such selection biases in order to capture the full range of CPS competence in a variety of different contexts and populations.

Finally, the assessment of CPS competence in general needs to be discussed. Typical CPS assessment tools as used in this study indicate whether a problem solver has acquired specific knowledge about a problem situation (i.e., the CPS process of knowledge acquisition) or whether a goal was reached (i.e., the CPS process of knowledge application). But strictly speaking, participants' behavior in the complex problem situation goes beyond these two core processes (see Dörner, 1986. In fact, in addition to the processes of knowledge acquisition and knowledge application, more specific processes are also considered important for solving complex problems. Some of these processes consist of engaging in strategic exploration to gather information, reducing and integrating information into a representation of knowledge, anticipating future developments and making plans, setting priorities and balancing goals, or evaluating and modifying problem solving behavior (e.g., Fischer et al., 2012; Funke, 2001; Greiff & Fischer, 2013; Wagener, 2001). Although there are already theoretical considerations about how to further develop the assessment of CPS to gain deeper insights into CPS competence (e.g., Greiff & Fischer, 2013), its practical implementation is still outstanding. In fact, more research on the evaluation of CPS performance beyond the two processes of knowledge acquisition and knowledge application is still needed (e.g., based on logfile analyses). Moreover, no direct measurement of GPSK was used in this study. That is, we only assumed that a better knowledge acquisition performance was based on a higher GPSK, but strictly speaking, we do not know whether this is truly the case. As a consequence, possible effects of the flexibility training used in the present study beyond the two CPS processes of knowledge acquisition and knowledge application (e.g., whether trainees had an increased GPSK, resulting in a larger number of tested hypotheses than participants without training) could not be evaluated. Future CPS training studies will therefore benefit considerably from the further development of CPS assessment tools that can capture additional and more specific CPS processes (e.g., Müller, Kretzschmar, & Greiff, 2013; Wüstenberg, Stadler, Hautamäki, & Greiff, 2014). Furthermore, in order to examine whether the increase in CPS competence was based on GPSK, future studies should (develop and) include a corresponding measure.

## Conclusion

As noted earlier, the trainability of CPS competence is an exciting issue and has become even more important since CPS competence was added to international educational large-scale assessments. However, it is not sufficient to theoretically assume its trainability and to give advice about how to improve CPS competence (e.g., OECD, 2014) if the empirical evidence is still missing. This study aimed to shed some light on the issue of the extent to which CPS is trainable. The improvement of CPS competence through experience with different problem situations might be possible. However, differential effects concerning the different CPS processes have to be considered. If CPS competence training were found to be effective only for the CPS process of knowledge acquisition (i.e., obtaining relevant information about an unknown problem) but not for the process of knowledge application (i.e., actually solving the problem), then it will be highly questionable whether CPS competence training would improve problem solving in real life (i.e., where the aim is to buy a ticket, not just to know about the functions of a ticket machine). Most important, further studies are needed to deepen our knowledge, especially in terms of the long-term effects of the training, the extent to which the training can be transferred to real-world problems, and the determinants of training success. In this respect, some crucial points for future research were highlighted in this study.

**Acknowledgements:** We thank Dietrich Wagener, Anette Kluge, as well as Mary M. Omodei for providing the software programs and Eigbert Riewald for his remarkable technical support. Furthermore, we thank Maik Böttcher, Samuel Greiff, Marcus Mund, and Cornelia Vogt for their helpful comments on earlier versions of this article.

**Declaration of conflicting interests:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be constructed as a potential conflict of interest.

**Author contributions:** The authors contributed equally to this work.

**Supplementary material:** The data are publicly available via the Open Science Framework and can be accessed at https://osf.io/n2jvy.



**Citation:** Kretzschmar, A. & Süß, H.-M. (2015). A study on the training of complex problem solving competence. *Journal of Dynamic Decision Making, 1*, 4. doi:10.11588/jddm.2015.1.15455

**Received:** 07 August 2015
**Accepted:** 03 December 2015
**Published:** 12 December 2015

## References

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., & Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. *Nature, 501*(7465), 97–101. doi:10.1038/nature12486

Anzai, Y., & Simon, H. A. (1979). The theory of learning by doing. *Psychological Review, 86*, 124–140. doi:10.1037/0033-295X.86.2.124

Bakken, B. E. (1993). *Learning and transfer of understanding in dynamic decision environments.* Massachusetts Institute of Technology.

Beckmann, J. F., & Goode, N. (2014). The benefit of being naïve and knowing it: the unfavourable impact of perceived context familiarity on learning in complex problem solving tasks. *Instructional Science, 42*(2), 271–290. doi:10.1007/s11251-013-9280-7

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video hames improve perception and cognition? *Frontiers in Psychology, 2*, 226. doi:10.3389/fpsyg.2011.00226

Bühner, M., Kröner, S., & Ziegler, M. (2008). Working memory, visual–spatial-intelligence and their relationship to problem-solving. *Intelligence, 36*, 672–680. doi:10.1016/j.intell.2008.03.008

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Dörner, D. (1976). *Problemlösen als Informationsverarbeitung* [Problem solving as information processing]. Stuttgart: Kohlhammer.

Dörner, D. (1986). Diagnostik der operativen Intelligenz [Assessment of operative intelligence]. *Diagnostica, 32,* 290–208.

Dörner, D. (1989). *Die Logik des Mißlingens: Strategisches Denken in komplexen Situationen* [The logic of failure: Recognizing and avoiding error in complex situations]. Reinbek bei Hamburg: Rowohlt.

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. doi:10.3758/BF03193146

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving, 4*(1), 19–41. doi:10.7771/1932-6246.1118

Fischer, A., Greiff, S., Wüstenberg, S., Fleischer, J., Buchwald, F., & Funke, J. (2015). Assessing analytic and interactive aspects of problem solving competency. *Learning and Individual Differences, 39,* 172–179. doi:10.1016/j.lindif.2015.02.008

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

Friedrich, H. F., & Mandl, H. (1992). Lern- und Denkstrategien - ein Problemaufriß [Learning and thinking strategies - A view on the problem]. In H. F. Friedrich & H. Mandl (Eds.), *Lern- und Denkstrategien* (pp. 3–54). Göttingen: Hogrefe.

Funke, J. (Ed.). (1999). Themenheft “Komplexes Problemlösen” [Special Issue: Complex Problem Solving]. *Psychologische Rundschau, 50*(4).

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: Kohlhammer.

Funke, J. (2006). Komplexes Problemlösen [Complex Problem Solving]. In J. Funke (Ed.), *Denken und Problemlösen* (pp. 375–446). Göttingen: Hogrefe.

Goettl, B. P., Yadrick, R. M., Connolly-Gomez, C., Regian, J. W., & Shebilske, W. L. (1996). Alternating task modules in isochronal distributed training of complex tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 38,* 330–346. doi:10.1177/001872089606380213

Gonzalez, C., Thomas, R. P., & Vanyukov, P. (2005). The relationships between cognitive ability and dynamic decision making. *Intelligence, 33*(2), 169–186. doi:10.1016/j.intell.2004.10.002

Goode, N., & Beckmann, J. F. (2010). You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. *Intelligence, 38*(3), 345–352. doi:10.1016/j.intell.2010.01.001

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537. doi:10.1038/nature01647

Greiff, S., & Fischer, A. (2013). Measuring complex problem solving: An educational application of psychological theories. *Journal for Educational Research Online, 5*(1), 38–58. Retrieved from: http://nbn-resolving.de/urn:nbn:de:0111-opus-80196

Greiff, S., Kretzschmar, A., Müller, J. C., Spinath, B., & Martin, R. (2014). The computer-based assessment of complex problem solving and how it is influenced by students' information and communication technology literacy. *Journal of Educational Psychology, 106*(3), 666–680. doi:10.1037/a0035426

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new assessment perspective. *Applied Psychological Measurement, 36,* 189–213. doi: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. doi:10.1037/a0031856

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61–83. doi:S0140525X0999152X

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test. Form 4* [Berlin Intelligence-Structure Test. Version 4]. Göttingen: Hogrefe.

Jensen, E. (2005). Learning and transfer from a simple dynamic system. *Scandinavian Journal of Psychology, 46,* 119–131. doi:10.1111/j.1467-9450.2005.00442.x

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

Kretzschmar, A., Neubert, J. C., & Greiff, S. (2014). Komplexes Problemlösen, schulfachliche Kompetenzen und ihre Relation zu Schulnoten [Complex problem solving, school competencies and their relation to school grades]. *Zeitschrift für Pädagogische Psychologie, 28*(4), 205–215. doi:10.1024/1010-0652/a000137

Kretzschmar, A., Neubert, J. C., Wüstenberg, S., & Greiff, S. (2016). Construct validity of complex problem solving: A comprehensive view on different facets of intelligence and school grades. *Intelligence, 54,* 55-69. doi:10.1016/j.intell.2015.11.004

Kulik, J. A., Kulik, C.-L. C., & Bangert, R. L. (1984). Effects of practice on aptitude and achievement test scores. *American Educational Research Journal, 21*(2), 435–447. doi:10.3102/00028312021002435

Meinz, E. J., & Hambrick, D. Z. (2010). Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: The role of working memory capacity. *Psychological Science, 21*(7), 914–919. doi:10.1177/0956797610373933

Müller, J. C., Kretzschmar, A., & Greiff, S. (2013). Exploring exploration: Inquiries into exploration behavior in complex problem solving assessment. In D'Mello, S. K., Calvo, R. A., and Olney, A. (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 336–337).

Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2014). Extending the assessment of complex problem solving to finite state automata: Embracing heterogeneity. *European Journal of Psychological Assessment, 31*(3), 181–194. doi:10.1027/1015-5759/a000224

Neubert, J. C., Mainert, J., Kretzschmar, A., & Greiff, S. (2015). The assessment of 21st century skills in industrial and organizational psychology: Complex and collaborative problem solving. *Industrial and Organizational Psychology, 8*(02), 238–268. doi:10.1017/iop.2015.14

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2002). The multiple facets of working memory - Storage, processing, supervision, and coordination. *Intelligence, 140,* 1–27. doi:10.1016/S0160-2896(02)00115-0

OECD. (2014). *Pisa 2012 results: Creative problem solving: Students' skills in tackling real-life problems* (Volume V). Paris: OECD Publishing.

Omodei, M. M., & Wearing, A. J. (1998). *Network Fire Chief.* La Trobe University.

Osman, M. (2010). Controlling uncertainty: A review of human behavior in complex dynamic environments. *Psychological Bulletin, 136,* 65–86. doi:10.1037/a0017815

Prensky, M. (2001). Digital natives, digital immigrants. Part I. *On the Horizon, 9*(5), 1–6. doi:10.1108/10748120110424816

Putz-Osterloh, W. (1987). Gibt es Experten für komplexe Probleme? [Are there experts for complex problems?]. *Zeitschrift Für Psychologie, 195,* 63–84.

Putz-Osterloh, W. (1988). Wissen und Problemlösen [Knowledge and problem solving]. In H. Mandl, H. Spada, & H. Aebli (Eds.), *Wissenspsychologie* (pp. 247–263). München: Psychologie Verlags Union.

Putz-Osterloh, W., & Lemme, M. (1987). Knowledge and its intelligent application to problem solving. *German Journal of Psychology, 11,* 286–303.

Saint-Mont, U. (2015). Randomization does not help much, comparability does. *PLoS ONE, 10*(7), e0132102. doi:10.1371/journal.pone.0132102

Sander, N. (2005). *Inhibitory and executive functions in cognitive psychology: An individual differences approach examining structure and overlap with working memory capacity and intelligence.* Aachen: Shaker.

Schaub, H., & Strohschneider, S. (1992). Die Auswirkungen unterschiedlicher Problemlöseerfahrungen auf den Umgang mit einem unbekannten komplexen Problem [The impact of expertise on the solving of an unknown, complex problem]. *Zeitschrift Für Arbeits- und Organisationspsychologie, 36,* 117–126.

Scherer, R., & Tiemann, R. (2012). Factors of problem-solving competency in a virtual chemistry environment: The role of metacognitive knowledge about strategies. *Computers & Education, 59,* 1199–1214. doi:10.1016/j.compedu.2012.05.020

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

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. doi:10.1037/e578442014-045

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. doi:10.1016/j.intell.2013.05.002

Stadler, M. J., Becker, N., Greiff, S., & Spinath, F. M. (2015). The complex route to success: Complex problem-solving skills in the prediction of university success. *Higher Education Research & Development,* 1–15. doi:10.1080/07294360.2015.1087387

Stern, E. (1993). Kognitives Training: Was verändert sich? Fragestellungen, Methoden und neuere Ergebnisse [Cognitve training: What changed? Issues, Methods and recent findings]. In L. Montada (Ed.), *Bericht über den 38. Kongreß der Deutschen Gesellschaft für Psychologie in Trier 1992* (Vol. 2, pp. 975–977). Göttingen: Hogrefe.

Strobach, T., Frensch, P. A., & Schubert, T. (2012). Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychologica, 140*(1), 13–24. doi:10.1016/j.actpsy.2012.02.001

Strohschneider, S. (1990). *Wissenserwerb und Handlungsregulation* [Knowledge Acquisition and Action Regulation]. Wiesbaden: Deutscher Universitäts-Verlag.

Süß, H.-M. (1996). *Intelligenz, Wissen und Problemlösen: Kognitive Voraussetzungen für erfolgreiches Handeln bei computersimulierten Problemen* [Intelligence, knowledge and problem solving: Cognitive prerequisites for successful behavior in computer-simulated problems]. Göttingen: 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., & Beauducel, A. (2015). Modeling the construct validity of the Berlin Intelligence Structure Model. *Estudos de Psicologia (Campinas), 32*(1), 13–25. doi:10.1590/0103-166X2015000100002

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy se and the acquisition of problem structure. *Cognitive Science, 20*(1), 75–100. doi:10.1207/s15516709cog2001_3

Vollmeyer, R., & Funke, J. (1999). Personen- und Aufgabenmerkmale beim komplexen Problemlösen [Person and task effects within complex problem solving]. *Psychologische Rundschau, 50,* 213–219.

Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios - Taxonomie, Entwicklung, Evaluation* [Psychological assessment with complex scenarios - taxonomy, development, evaluation]. Lengerich: Pabst.

Wagener, D. (2007). *START-C. Testbatterie für Berufseinsteiger - Computer* [Test battery for young professionals — computers]. Göttingen: Hogrefe.

Wagener, D., & Conrad, W. (2001). FSYS 2.0. Testmanual. University of Mannheim.

Wagener, D., & Wittmann, W. W. (2002). Personalarbeit mit dem komplexen Szenario FSYS [Human resource management using the complex scenario FSYS]. *Zeitschrift Für Personalpsychologie, 1,* 80–93.

Walberg, H. J., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal, 20*(3), 359–373. doi:10.3102/00028312020003359

Wallach, D. (1998). *Komplexe Regelungsprozesse: Eine kognitionswissenschaftliche Analyse* [Complex control processes: A cognitive analysis]. Wiesbaden: Deutscher Universitäts-Verlag.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle, WA: Hogrefe.

Weinert, F. E., & Waldmann, M. R. (1988). Wissensentwicklung und Wissenserwerb [Development and acquisition of knowledge]. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 161–202). München: Psychologie Verlags Union.

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. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: Process, trait and content determinants* (pp. 77–104). Washington, D.C.: APA.

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., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The Role of strategy knowledge for the application of strategies in complex problem solving tasks. *Technology, Knowledge and Learning, 19*(1), 127-146. doi:10.1007/s10758-014-9222-8