Doctoral thesis submitted to
the Faculty of Behavioural and Cultural Studies
Heidelberg University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy (Dr. phil.)
in Psychology

Title of the publication-based thesis

Complex Problem Solving and Organizational Psychology: Of Nomological Networks, Microworlds, and Interdisciplinary Dialogue

presented by
Dipl.-Psych. Jonas Christian Neubert

year of submission
2016

Dean: Prof. Dr. Birgit Spinath
Advisor: Prof. Dr. Dr. h.c. Joachim Funke
Prof. Dr. Christiane Hipp

Table of content

List of scientific publications regarding the publication-based thesis

I. Publication

Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2015). Extending the Assessment of Complex Problem Solving to Finite State Automata: Embracing Heterogeneity. *European Journal of Psychological Assessment, 31,* 181-194.

II. Publication

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*(1), 55–69.

III. Publication

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*, 238–268.

Additional publications

Fischer, A., & Neubert, J. C. (2015). The multiple faces of complex problems: A model of complex problem solving competency and its assessment. *Journal of Dynamic Decision Making*, *1*(6), 1–14.

Greiff, S., Kretzschmar, A., Müller, J. C., Spinath, B., & Martin, R. (2014). Computer-based assessment of Complex Problem Solving in educational contexts and how it is influenced by students' level of Information and Communication Technology literacy. *Journal of Educational Psychology*, *106*(3), 666–680.

Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, *36*, 37–48.

Greiff, S., Neubert, J. C., Niepel, C., & Ederer, P. (2015). Problem solving—facilitating the utilization of a concept towards lifelong education. *International Journal of Lifelong Education*, *34*(4), 373–375.

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.

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.

Neubert, J. C., Lans, T., Mustafic, M., Greiff, S., & Ederer, P. (2017). Complex Problem-Solving in a Changing World: Bridging Domain-Specific and Transversal Competence Demands in Vocational Education. In M. Mulder (Ed.), *Competence-based Vocational and Professional Education* (pp. 953–969). Cham: Springer.

“There are things known and
there are things unknown, and
in between are the doors of
perception.”

Aldous Huxley

1. Introduction

The founding fathers of Complex Problem Solving (CPS) research utilized the advent of computer-technology to connect problem solving research to the affordances of everyday life: By building on computer-simulated models of complex problem situations, such as the handling of a developmental aid program or the management of a small town, they were able to connect the advantages of laboratory research to problem solving beyond simple and static tasks (e.g., Brehmer & Dörner, 1993). During the last decades, this tradition has led to a thriving field of CPS research with interesting implications for a number of applied settings (see Frensch & Funke, 1995a; J. Funke, 2006 for overviews).[1] As a result, problem solving assessment instruments have been included in large-scale assessment efforts in the educational domain, such as the Programme for the International Student Assessment (PISA, OECD, 2014) and successfully handling complex problems is included in nearly every list of important prerequisites for the 21st century (e.g., P. Griffin, McGaw, & Care, 2012; National Research Council, 2012; OECD, 2013a; World Economic Forum, 2016).

While recent research efforts have sparked interest in the construct and it's assessment in the educational domain, the goal of this thesis is (re-)connecting CPS research to another area of application: Organizational Psychology (OP),[2] where these efforts have been greeted with much less discussion and exchange. To establish an interaction with this domain of research and practice, several aspects within current CPS research need to be considered. To this end, the first core paper of the thesis explores the relation of CPS to one of the most important psychological constructs in OP, namely intelligence, thereby complementing the nomological network of the construct (Chapter 2). The second core paper of the thesis is closer aligned with the practical side of utilizing a construct in OP and considers the valid and reliable assessment of CPS with the help of finite state automata (Chapter 3).

---

[1] Please note that there is a number of synonyms for CPS being used in the literature, including Dynamic Decision Making (e.g., Gonzalez, Lerch, & Lebiere, 2003; Gonzalez, Vanyukov, & Martin, 2005), Dynamic Problem Solving (Greiff, Wüstenberg, & Funke, 2012), and Interactive Problem Solving (e.g., OECD, 2013b). Some of the labels focus the attention on subtly different aspects of the process of dealing with complex problems (e.g., decision making vs. problem solving) but their core seems similar: A focus on human interaction with problems characterized by complexity and dynamics (see Fischer, Greiff, & Funke, 2012; Fischer, Holt, & Funke, 2015).

[2] Using the shorter „organizational psychology" rather than industrial (and) organizational psychology seems to be a trend in the field which is followed here (e.g., Ryan & Ford, 2010).

The third core paper of the thesis builds on this foundation in terms of construct and assessment and explores the connections between CPS and OP in a discussion of researchers and practitioners from both fields (Chapter 4). In summary, the thesis aims at complementing the current state of insights within CPS research in light of an application in OP and the start of an exchange between both domains. But before digging deeper into potential routes for connecting CPS and OP, a closer look at the research tradition of CPS is warranted.

1.1. Historical roots and earlier problem solving research

One of the points of departure for the initiation of CPS research was the perceived mismatch between ‘classical’ laboratory research on human problem solving and the everyday complex problem environments (e.g., Dörner, 1989b; Dörner & Reither, 1978). More specifically, ‘classical’ problem solving research during the mid-70s focused on well-defined problems, such as the Tower of Hanoi to investigate a range of features of human problem solving, such as the transfer of solutions to new problems (e.g., Kotovsky, Hayes, & Simon, 1985; Newell & Simon, 1972; see for example Dunbar, 1998; Mayer, 2011, for more general overviews).

Interestingly, the underlying definitions of a problem and problem solving have been shown to be quite stable during the years despite various shifts of attention: Contemporary reviews and research articles still refer to Gestalt psychologist Karl Duncker and his definition of a problem as “when a living creature has a goal but does not know how this goal is to be reached.” (Duncker, 1945, p. 1; for example referred to in Fischer, Greiff, & Funke, 2012; Mayer, 2011). Similarly, the definition of problem solving as the process of searching for operators to transfer a given state into a goal state (as introduced by Newell & Simon, 1972) has remained important for current (complex) problem solving research (and with it the notion of the problem space; see for example Dunbar, 1998; Fischer et al., 2012). Even more, these attempts at describing the nature of human problem solving efforts have roots going back to the very beginning of (modern) psychological research and the works of psychologists from the Würzburg School probing into human thinking via experiments and self-reports at the end of the 19th century (e.g., J. Funke, 2006; Rollett, 2008; ter Hark, 2010).

The Tower of Hanoi and similar problems used in ‘classical’ problem solving research (e.g., the missionary-cannibals problem, see for example Reed, Ernst, &

Banerji, 1974) can be formally described in an exhaustive way (e.g., via state-transition diagrams), facilitating the development of concepts such as the problem and solution space (Newell & Simon, 1972; Nilsson, 1971). The formal description, in turn, allows for the comparison of participants' problem solving efforts in different isomorphic or homomorphic problems with the same or similar underlying structure but differences in problem presentation or description (e.g., Reed et al., 1974; Simon & Hayes, 1976).

The possibilities of formal description consequently led to an impressive amount of insights, for example regarding the influence of different problem instructions or differences in problem solving strategies (for overviews see e.g., Kotovsky et al., 1985; Mayer, 2011). In a study by Herbert Simon (1975), the author was able to distinguish and formally describe several possible strategies of solving the Tower of Hanoi problem, such as a recursive decomposition of the problem into smaller sub-problems, the application of simple movement rules, or the recursion to more "rote" procedures by storing correct solution patterns in memory. This formal description in turn, allowed for the construction of programs able to solve the Tower of Hanoi problem autonomously, progressing towards the goal of general information-processing models of problem solving and the so-called "general problem solver" built to solve all kinds of different problems (Newell & Simon, 1972; see Ohlsson, 2012, for a critical account).

Even though the framework of looking into the problem solving efforts of participants on comparably simple problems such as the Tower of Hanoi led to impressive insights such as the ones presented above, there are also a number of problems with the framework that resulted in modified approaches to human problem solving (see also Ohlsson, 2012; VanLehn, 1989). The research tradition of CPS therefore departed from the research on small and formally well-describable problems to investigate human interaction with more complex problems.[3]

[3] Other examples of extending the range of problem solving research include the consideration of external representations outside the problem solver's mind for the process of solving problems (e.g., Zhang & Norman, 1994) or the investigation of the role of expert knowledge when perceiving, representing, and solving problems in specific domains, highlighting, for example, considerable differences in solution strategies (e.g., Larkin, McDermott, Simon, & Simon, 1980).

1.2. The advent of CPS research

More specifically, Dietrich Dörner and colleagues (Dörner, 1989b; Dörner & Reither, 1978) started from the stark contrast between the problems utilized in problem solving research, such as the Tower of Hanoi and real-world problems, such as the ones arising when managing a company, fighting a blazing fire or trying to improve the living conditions in a country via developmental aid. Aided by the availability of (comparably cheaper) computer technology, these authors attempted to combine the benefits of studying problem solving behavior in the laboratory (e.g., experimental control of the problem situations) with the features of complex everyday problems (e.g., Brehmer & Dörner, 1993).

To this end, the pioneers of CPS research employed computer-simulated problems, so-called 'microworlds', to simulate complex problem environments, such as the management of a small town (LOHHAUSEN, Dörner, Kreuzig, Reither, & Stäudel, 1983), developmental aid programs (MORO, Strohschneider, 1986; Strohschneider & Güss, 1999), or taking care of a company (TAILORSHOP, Dörner, 1979b; J. Funke, 1983; Putz-Osterloh, 1981; see Gonzalez, Vanyukov, & Martin, 2005; J. Funke, 2003, for overviews on different microworlds). The computer-simulated nature of the problems allowed researchers to investigate human interaction with complex problem environments within the controlled environment of a laboratory (i.e., being able to observe different participants' problem solving efforts under controlled circumstances), while the nature of the problems under study was changed towards features, such as complexity, dynamics, and intransparency (Brehmer & Dörner, 1993, see also Chapter 2).

For example, the microworld TAILORSHOP is simulating the management of a small company producing shirts (Dörner, 1979b; J. Funke, 1983). Initially programmed on the pocket calculator of Dietrich Dörner, the TAILORSHOP microworld has shown remarkable longevity to this day and has been utilized in numerous studies on CPS (e.g., Barth & Funke, 2010; Danner et al., 2011; Öllinger, Hammon, Grundherr, & Funke, 2015; Süß, 1996; Wittmann & Hattrup, 2004). Within the TAILORSHOP simulation, problem solvers have to control the fate of a small company by choosing appropriate levels for workers' wages, maintenance of machinery, prices of the shirts and other parameters and thereby maximize the company's value and profit over the course of several simulated years (see J. Funke, 1983, 2010; Süß, 1996 for more details). The problem solver therefore has to explore and decide on a range of different factors, explore their interrelation with several other variables (partially not directly visible), and monitor the effect of his or her interventions (see Chapter 2 and e.g., Fischer et al., 2012; J. Funke, 2010; Süß, 1996). Through their simulation of problem situations exhibiting several features deemed to be important in real-life problem solving in complex environments, microworlds such as the TAILORSHOP therefore allowed for the development of a research tradition focusing on complex problem solving (see Chapter 2 and 3 as well as Frensch & Funke, 1995a; J. Funke, 2003, for broader overviews on the construct of CPS and its' assessment).

The introduction of microworlds simulated on computers led to a range of interesting findings in the tradition of CPS research. Dörner observed participants interaction with computer-simulated microworlds and identified a number of errors human problem solvers typically exhibit when dealing with complex problems (Dörner, 1989b, 1990). For example, the insufficient elaboration of goals leads to problem solving behavior following a "repair service policy" (Dörner, 1990, p. 19), where obvious errors and problems are dealt with at the expense of problematic long term developments that are only recognized once it is too late. Similarly, a lack of hypotheses about the system and problems in understanding the temporal development of the system lead to problem solving behavior that neglects long-term and side effects of the system and a failure to account for dynamic changes in the system (see e.g., J. Funke, 2003, for a comprehensive account of Dörner's findings).

Importantly for the quest of this dissertation and the link between CPS research and OP, the findings and approach of early problem solving research as well as the investigations by Dörner and colleagues were taken up by researchers and practitioners concerned with a range of research interests, including those in the domain of OP.[4] For example, in addition to his pioneering work in investigating human problem solving (e.g., the introduction of the dual-space theory, see above), Herbert Simon also became a prominent figure in the research on management and organization, writing some of the most influential books of the discipline (March & Simon, 1993; Simon, 1997)[5] and connecting his view on problem solving with the context of organizations (e.g., Simon, 1979).

---

[4] According to Google Scholar, the book by Dietrich Dörner on the 'Logic of failure' has been cited over 2000 times since its publication in 1989 and the book is currently available in its 13th edition.
[5] Both books ranked among the top ten in a list of most influential management books of the 20th century, rated by scholars of the discipline (Bedeian & Wren, 2002).

More closely aligned with CPS and the work of Dörner, the computer simulations developed for the investigation of CPS, for example TAILORSHOP (see above), were also applied for the purpose of selecting employees and the findings of Dörner were taken up in OP as well (see Chapters 3 and 4). Intriguingly enough, recent advances in CPS research, such as the introduction of multiple complex systems in assessment, have not been taken up in OP. But before digging deeper into the relation between CPS and Organizational Psychology, a closer look at the construct and its assessment is necessary.

2. The construct of CPS and its nomological network

2.1. The construct of CPS

2.1.1. Defining CPS

The growing research on human interaction with complex problems since the field's inception has resulted in a number of definitions of complex problems and complex problem solving. Conveniently, Peter Frensch and Joachim Funke assembled a comprehensive overview of different definitions and their respective focus on various elements, and developed a summarizing definition of CPS:

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

As can be easily seen, the definition includes the traditional definition of problem solving via given states and goal states (see Chapter 1), as well as the notion of barriers between given states and goal states as introduced by Dörner (e.g., Dörner, 1979a; J. Funke, 2003). More interesting for the notion of *complex* problem solving is the part that follows: The given state, goal state, and barriers are said to exhibit specific features, namely (a) novelty (b) complexity, (c) dynamics, and (d) intransparency (see e.g., Brehmer, 1992, for a different set of features). Interestingly, the focus on specific features of complex problems investigated under the label of CPS has undergone minor changes; recent investigations have largely dropped the explicit consideration of the feature of novelty and put more emphasis on additional features, most notably interrelatedness and politely (i.e., multiple goals, see J. Funke, 2003; J. Funke, Fischer, & Holt, in preparation, for a broader discussion of problem features).[6] In line with these developments, the focus of this dissertation is on complex problems characterized by the features of complexity, connectivity, dynamics, intransparency, and politely (see e.g., J. Funke, 2001, 2003, for encompassing discussions of the features).

[6] Interestingly, the differentiation of the features being related to the given state, the goal state or the barriers has not received specific attention in favor of a more general description relating the features to the problem as a whole (but see e.g., Fischer et al., 2012).

In the research tradition of CPS, a problem is considered *complex* if it contains multiple elements that require consideration (e.g., Quesada, Kintsch, & Gomez, 2005; researchers from other disciplines have referred to other characteristics, such as “regularities”, when writing about complexity, e.g., Gell-Mann, 1995). For example, managing a small company in the TAILORSHOP simulation (and in reality for that matter) requires decisions on a range of different factors, such as workers’ wages, maintenance of machinery, prices of shirts and other elements. On the side of the problem solver, the features of complex problems lead to very different requirements (see e.g., Fischer et al., 2012; J. Funke, 2003 for overviews). In the case of complexity, the necessity arises to cope with the multitude of relevant elements, for example through a focus on a subset of factors in the TAILORSHOP simulation (Fischer et al., 2012). With regard to its formal description, the notion of complexity has been the subject of discussions inside and outside of CPS research without reaching a definite measure of complexity. For example, Quesada et al. (2005) discuss different indicators of complexity including the time an algorithm needs to solve a given problem, the number of relations that need to be considered simultaneously, and the size of the problem space, each with their own advantages and problems (see e.g., J. Funke, 2003; Quesada et al., 2005 for overviews).

The feature of *connectivity* describes the interrelations of elements that need to be handled in complex problems (J. Funke, 2003). For example, in the TAILORSHOP example, the number of workers is influencing the total cost of personnel, which in turn is influencing both, worker satisfaction and company profit. These interrelations need to be considered when trying to increase worker satisfaction while also being responsible for company profit (i.e., considering side-effects of one’s actions, Dörner, 1990). On the side of the problem solver, the feature of connectivity leads to the need for model building, that is, finding and adequate representation of the problem’s internal structure to predict intervention’s effects, for example when increasing the number of workers (e.g., J. Funke, 2003). Interestingly, the number of connections between a problem’s elements has been also proposed to indicate its complexity (i.e., relational complexity, Quesada et al., 2005), indicating problems of differentiating both features.

Another feature of complex problems is focused on the passage of time and the related changes in problems: *Dynamics*. That is, the exemplary management of a TAILORSHOP simulation requires the reaction to changes in the problem’s elements over time or in response to the problem solvers efforts (e.g., a decrease in customer demand). Due to the dynamics centrality in differentiating CPS from other domains of problem solving, CPS is also sometimes referred to as Dynamic or Interactive Problem Solving (e.g., Fischer, Greiff, et al., 2015) and a whole research strand is dedicated to the exploration of decision making in dynamic environments and the resulting requirements (e.g., Brehmer, 1992; Gonzalez et al., 2005; Rouwette, Größler, & Vennix, 2004). Similar to complexity and connectivity, dynamics-related differentiations of complex problems have been also proposed for taxonomies of CPS, for example differentiating problems according to their time invariant or dynamic nature, the continuous vs. discrete progression of time, and the degree of time pressure (Quesada et al., 2005).

The problem feature of *intransparency* is describing the obfuscated nature of the problem's internal organization and interrelations and thus, the requirement to explore the problems underlying rules and relations actively. In the TAILORSHOP, the problem solver does not know about the internal organization of the microworld's variables and therefore has to actively explore the interrelations from their reaction to input manipulations and the corresponding outputs or guess them from prior knowledge (i.e., information generation). According to Dörner and Wearing (1995), the intransparency of real-life decisions with regard to the situation, the consequences of actions, and the prerequisites of decisions is meant to be recreated in microworlds used for the study of CPS (Dörner & Wearing, 1995; see Rigas, Carling, & Brehmer, 2002, for a more general argument in the same direction of reproducing real-life situations in the laboratory).

Finally, the feature of *polytely* is referring to the existence and requirement of addressing the complex problem situation in light of multiple potentially conflicting goals. While the problem solver has to strive for the overarching goal of maximizing the company profit in the TAILORSHOP microworld, a lot of different sub-goals have to be balanced to achieve this. For example, the limitations in financial resources require decisions with regard to the investment of worker's wages and machinery maintenance, both representing different sub-goals on the road to a profitable company (i.e., satisfying levels of worker satisfaction and breakdowns of machinery).

The translation of these problem features into computer simulations has resulted in debates on the adequate operationalization of problems and their features (J. Funke, 2014a; Greiff & Martin, 2014; Quesada et al., 2005; Schoppek & Fischer, 2015, see Chapter 3). Specifically, it remains unclear, whether different problems simulated on the computer actually represent the same underlying (latent) construct (J. Funke, 2010, 2014a), whether and in what way different simulations require (qualitatively) different cognitive processes (Fischer, 2015; J. Funke, 2014a; Schoppek & Fischer, 2015), and whether one can adapt an overarching framework of CPS to accommodate the differences in problem features and requirements (e.g., Fischer & Neubert, 2015; J. Funke et al., in preparation). We will come back to the topic of translating the construct of CPS into measurement instruments in Chapter 3.

At least as important for the goal of this thesis, the connection of CPS to organizational psychology, is that the notion of complexity and complex problems or tasks has also found its way into the domain of organizational psychology (e.g., Axley & McMahon, 2006; Tsoukas & Hatch, 2001), for example through concepts such as task complexity in the work analysis literature (Campbell, 1988; Hackman, 1969; Jenkins, 2009; Wood, 1986; see Hærem, Pentland, & Miller, 2015, for a recent extension). Independently from CPS research, it has led to similar problems of finding an appropriate measure of complexity and debates on important features (e.g., Hærem et al., 2015; Liu & Li, 2012), offering potential routes for interdisciplinary cooperation (see Chapter 4 and 5).

2.1.2. The process of solving complex problems: Knowledge acquisition and knowledge application

The development of CPS from traditional problem solving research has influenced the differentiations in theoretical and empirical investigations when investigating the process of human interaction with complex problems (see e.g., Fischer et al., 2012, for a discussion). For example, the general differentiation between processes related to the understanding and representation of the problem or the acquisition of knowledge from those related to the search and monitoring processes when trying to solve the problem or the application of knowledge (e.g., Mayer, 2011; Novick & Bassok, 2005; VanLehn, 1989) has found its way from the traditional literature on human problem solving into current research on CPS (e.g., Beckmann & Goode, 2013; Fischer et al., 2012; J. Funke, 2001; Greiff, Wüstenberg, & Funke, 2012; Neubert, Kretzschmar, Wüstenberg, & Greiff, 2015).

In CPS research, knowledge acquisition is the process through which individual problem solvers accumulate the knowledge to (more or less) successfully handle a problem characterized by the features presented above (i.e., establishing an adequate representation of the problem, Mayer, 2011; Novick & Bassok, 2005). Fischer et al. (2012) present a process model of complex problem solving that integrates different factors influencing the process of acquiring knowledge in CPS. In their model, knowledge acquisition is characterized by (a) an exploration of the problem situation with known and ecologically rational strategies leading to (b) knowledge about a system's elements and reactions to inputs as well as internal relations and structure, while (c) working under the restrictions of the human mind and hence, the need to reduce information (see also Jonassen, 2004, for a similar process description from the perspective of teaching problem solving). Among other elements, the model is highlighting the role of systematic strategy use (Fischer et al., 2012; see also Wüstenberg, Stadler, Hautamäki, & Greiff, 2014) and differences in knowledge on a problem's instances and a underlying structure (i.e., input-output or instance-based knowledge vs. structural knowledge, Schoppek, 2002; see also Gonzalez, Lerch, & Lebiere, 2003).

There is a host of research from the contexts of problem solving, causal reasoning and conceptual change, that offers plenty of insights on the process of knowledge acquisition in CPS that is also relevant for applications in organizational psychology. Of special interest from the perspective of CPS are the difficulties arising for knowledge acquisition in more complex problem settings (e.g., Dörner, 1980, 1989b; Kuhn et al., 1995). For example, drawing causal inferences becomes more difficult, when the number of potentially relevant influences increases (see the feature of complexity, above) and in the absence of opportunities for controlled comparisons. In the exemplary case of the company producing shirts simulated in the TAILORSHOP microworld, the effect of investments into the maintenance of machinery can be hardly explored in isolation. Instead, the problem solver has to infer the variable's relevance from the broader behavior of the microworld during the simulated timeframe (Kuhn et al., 1995). Similarly, problem solvers in the business context have to rely on inferences based on covariation or so-called generalized inclusion inferences instead of controlled experiments (see Kuhn et al., 1995) or use strategies and heuristics tailored to the specific situation (Gigerenzer & Gaissmaier, 2011). These strategies and heuristics in complex problem situations in turn, should be of interest for organizational psychology (e.g., during assessment, see Neubert, Mainert, Kretzschmar, & Greiff, 2015, and Chapter 4).

The second phase of problem solving is concerned with the solution to the problem and includes the cognitive operations of planning, executing, and monitoring (e.g., Mayer, 2011). In Fischer et al. (2012)'s process model of CPS, the phase of knowledge application is characterized by (a) the need to predict the problem's development and the consequences of one's own actions and (b) monitoring processes dedicated to the tracking of progress and potential misrepresentations of the problem. The latter part of the monitoring processes, the control of the problem's representation during problem solving already points towards the intertwined nature of both knowledge acquisition and knowledge application. Similar to the case of knowledge acquisition, the features of complex problems lead to issues in the knowledge application phase of CPS. For example, Dörner (1989b) already pointed towards typical problems of insufficient planning and the lack of correcting errors when dealing with complex problem situations. That is, the simulated managers of the shirt factory tend to make plans for the future building on problematic expectations (e.g., underestimating the effect of their interventions, overlooking mistakes they made). The examples of solutions riddled by unexpected problems and subsequent mishandling in the domain of OP are numerous (see Law & Callon, 2006; Volmerg, Leithäuser, Neuberger, Ortmann, & Sievers, 1995, for examples) and show the relevance of handling complex problems in organizational settings.

Interestingly, there are already promising links to research in the business context on this basic level of problem solving processes. For example, there is a whole research tradition concerned with developing appropriate (mathematical) representations and subsequently searching for solutions for problems in the domain of operations research that could serve as a point of departure for interdisciplinary collaboration and exchange (e.g., planning and optimizing supply chains for car manufacturers or the scheduling of airlines, Churchman, Ackoff, & Arnoff, 1971; Hillier & Lieberman, 2005; R. Klein & Scholl, 2004). On the other hand, research in organization science has also emphasized the general boundaries of rational problem solving when understanding and especially handling complex problems and detailed the influence of probabilistic and random processes, again offering plenty of opportunities for conceptual and empirical exchange (e.g., the famous garbage can model of organizing, M. D. Cohen, March, & Olsen, 1972; see also Elster, 1987; Gavetti, Levinthal, & Rivkin, 2005; Nickerson & Zenger, 2004; Rasmussen, 1997; Rouwette et al., 2004; Sterman, 2006).

In CPS, the differentiation of different factors influencing the problem solving processes has recently led to the development of encompassing models describing the competencies necessary to deal with complex problems (e.g., Fischer & Neubert, 2015; J. Funke et al., in preparation). For example, in the model proposed by Fischer and Neubert (2015), the components of knowledge, skills, abilities, and other factors are brought together, to form a comprehensive picture of the different factors relevant in solving specific complex problems. Interestingly, the authors thereby refer to a well-established differentiation from the area of organizational psychology, originating in the analyses of work tasks and the resulting requirements (e.g., Campion et al., 2011; Fischer & Neubert, 2015, see Chapter 5).

2.2. The nomological network of CPS

The multitude of factors included in competency models such as the one proposed by Fischer and Neubert (2015) highlights the breadth of constructs that are potentially relevant in dealing with complex problems (see Kaslow et al., 2007; Sanchez & Levine, 2009, for a broader perspective on competency-based models). Thus, the nomological network of CPS, its relation to other constructs becomes an important target of theoretical and empirical inquiry.

Generally speaking, the “nomological network defining the theory consists of the interpreted axiomatic system plus all of the empirical laws derived from it.” (Kane, 2001, p. 321). That is, the theoretical constructs are connected to each other and to observable variables and these variables in turn can be subjected to empirical examination (Kane, 2001). So without delving too deep into the debate on different ways of validating constructs (see e.g., Chapelle, Enright, & Jamieson, 2010; Kane, 2001, 2006), to qualify for a distinct construct, successfully handling complex problems needs to be sufficiently distinct from indicators of personality dimensions, business knowledge, or intelligence (see already Cronbach & Meehl, 1955; and more recently the Standards for Educational and Psychological Testing by AERA, APA, & NCME, 1999).[7]

---

[7] An aspect of validity theory that has received little attention in CPS research so far is the need to validate assessment for specific purposes of application (Kane, 2001). Current approaches to validity operate on the basis of validating assessment approaches in light of specific areas of application (e.g., Chapelle, Enright, & Jamieson, 2010). In contrast, empirical investigations into the validity of CPS research have focused on the interplay of constructs and measurement instruments without taking the

If tests of general intelligence or other already established constructs already include all the necessary knowledge to be gathered on a person's capabilities in handling complex problems, there is nothing to be gained from an additional construct of CPS or a dedicated assessment of CPS for that matter (Kretzschmar, 2015; Kretzschmar, Neubert, Wüstenberg, & Greiff, 2016; Süß, 1996; see also Oh, 2015). For the connection of CPS to organizational psychology, a conceptually and empirically stable nomological network is therefore immensely important, as the search for commonalities and distinctions needs to be built on a solid foundation on the side of both domains of research. To this end, a range of (empirical) investigations have highlighted the relation of CPS performance to a variety of psychological constructs.

The number of potentially relevant constructs is extremely large, as the interplay of (non-)cognitive processes involved in handling complex problems can be considered complex, too (Fischer & Neubert, 2015; J. Funke, 2010; but see Gigerenzer & Gaissmaier, 2011, for a contrary position). Encompassing models of human intellect, such as the PPIK theory developed by Ackerman (1996) or competency models of CPS (Fischer & Neubert, 2015; OECD, 2013b) might therefore serve as a point of departure for broader investigations into the nomological network of CPS. Fortunately for this thesis, previous research has already assembled a range of insights on the nomological network of CPS, clarifying some of the relations to other constructs (but see Footnote 7). In light of the multitude of potentially relevant factors, specific examples will therefore serve the purpose of highlighting the general approach to the validation of CPS assessment, before a specific case with high relevance for organizational psychology, namely that of intelligence, is taken into closer examination (see Chapter 2.3).

To give an example of the general approach to validating CPS assessment, a study by Greiff and Neubert (2014) will be examined in the following, which shed light on the relation of complex problem solving efforts to measures of personality and reasoning ability. Both factors, personality and reasoning ability have been shown to be important prerequisites for occupational performance across the lifespan rendering them important parts of the nomological network of CPS for the purpose of this dissertation (e.g., Judge, Higgins, Thoresen, & Barrick, 1999, see below).

---

purpose of assessment into account (e.g., highlighting interindividual differences vs. competency levels).

Specifically, Greiff and Neubert (2014) investigated the relation of personality and reasoning ability to the performance in so-called MICRODYN tasks, said to capture the CPS skills of individuals (Greiff et al., 2012, see Chapter 3 on the assessment of CPS). In their study conducted with high school students in Germany, CPS performance was related to personality as conceptualized in the five-factor model of personality (Costa & McCrae, 1992) and measured by a short version of the NEO personality inventory (NEO-GER, Rammstedt & John, 2007) and reasoning ability as measured with the help of the Culture Fair Test (CFT 20-R, Weiß, 2006, see Chapter 2.3 for a closer look at the relation of CPS and intelligence).

Encompassing models of human problem solving, such as the PPIK theory developed by Ackerman (1996) or the competency model of CPS by Fischer and Neubert (2015) include personality and interests as important factors when dealing with complex problems, but contrary to these predictions, the personality factors as indicated by the NEO-GER were related to CPS performance in partially unexpected ways (see Greiff & Neubert, 2014, for more details). In contrast to it's relation to performance indicators in the world of work, Conscientiousness was actually *negatively* related to knowledge acquisition and knowledge application, which was not expected before ($\beta = -.10$; $p < .05$, $\beta = -.10$; $p < .05$). Individuals ranking higher on the indicator of Conscientiousness, that is supposed to indicate facets, such as self-discipline and dutifulness, were actually performing worse in dealing with complex problems (see e.g., Judge et al., 1999, for a contrasting picture for job performance).

More in line with expectations, Neuroticism was negatively related to both dimensions of CPS ($\beta = -.12$; $p < .05$, $\beta = -.14$; $p < .05$), and Agreeableness was also surprisingly negatively related to CPS performance ($\beta = -.21$; $p < .05$, $\beta = -.18$; $p < .05$). In contrast to the surprising associations of the dimensions reported above, Extraversion and Openness were *not* significantly related to CPS performance (all $p > .10$), although one would have expected a positive relation for both dimensions of personality. Although potentially also due to methodological problems of the instrument targeting personality,[8] the findings of Greiff and Neubert (2014) are puzzling indicators of the complex relation between personality and performance in complex problem solving. On the other hand and closer related to the question of a distinct construct of CPS, the findings also highlight the differences between CPS

[8] The short scale of personality exhibited sub-optimal psychometric features (i.e., internal consistencies ranging between $\alpha = .52$ and .72).

performance and measures of personality, thereby supporting the construct validity of CPS (but see Footnotes 7 and 8) as well as the need for proper assessment instruments targeting all of the involved constructs (see Chapter 3 for more details on instruments targeting CPS).

From the perspective of organizational psychology and the overarching goal of this thesis, the (lack of a coherent) relation between indicators of personality and measures of performance indicating CPS is indeed important. Systematic reviews of previous (empirical) research and meta-analyses have sharpened the picture of the influence of personality in the organizational setting (e.g., Barrick & Mount, 1991; Judge, Bono, Ilies, & Gerhardt, 2002; Judge et al., 1999; Kohn & Schooler, 1982; Lord, de Vader, & Alliger, 1986; Poropat, 2009). For example, Judge, Bono, Ilies, and Gerhardt (2002), examined the influence of personality traits on leadership outcomes in a systematic review and meta-analysis of previous research. They found a significant and positive relation between Extraversion, Conscientiousness, and Openness and two indicators of leadership (i.e., leadership emergence and leadership effectiveness, $\rho = .24$ to $.31$). All of these relations contrast with the negative or zero-relation between the respective dimension of personality and CPS performance in Greiff and Neubert (2014). Especially for Conscientiousness this finding is troubling, as the personality dimension is considered important for occupational success even beyond general mental ability (Schmidt & Hunter, 2004) but exhibited a *negative* relation to CPS performance ($\beta = -.10$ for knowledge acquisition and $\beta = -.10$ for knowledge application; both $p < .05$).

Similarly, the relation between Agreeableness and leadership was weak but positive ($\rho = .08$) in Judge et al.'s (2002) meta-analysis, but negative and significant for the personality dimension and CPS performance ($\beta = -.21$ and $-.18$; both $p < .05$) in Greiff and Neubert (2014). The only overlap in findings exist for Neuroticism, where both, the meta-analysis on leadership ($\rho = -.24$) and the relation to CPS performance ($\beta = -.12$ and $-.14$; both $p < .05$) indicate a significant negative relation. In light of the complex nature of leadership and the requirement of productively dealing with complexity there (e.g., Vargas Cortes & Beruvides, 1996), these findings of differing relations between personality, leadership, and CPS performance are indeed puzzling and invite further empirical research (see Chapters 4 and 5). This clarification of relations is especially important as the conceptual requirements in leadership and CPS overlap to at least some extent (Neubert, Mainert, et al., 2015, see Chapter 4).

The other construct included in Greiff and Neubert (2014), reasoning ability, points to another important construct, namely intelligence. In the study by Greiff and Neubert (2014), reasoning ability was moderately to strongly related to performance in the CPS assessment: The CFT predicted both knowledge acquisition ($\beta = .62$; $p < .05$; $R^2 = .38$) and knowledge application in MICRODYN ($\beta = .51$; $p < .05$; $R^2 = .26$) to a substantial degree. That is, interindividual differences in the performance in the CPS test could be substantially predicted[9] by the corresponding differences in the different elements of the CFT (i.e., numerical series, classifications, matrices, and topological reasoning). On the one hand this strong relation comes to no surprise, as problem solving has been prominently featured in nearly all definitions of intelligence (e.g., Raven, 2000). Nonetheless or even more so, the relation between the two constructs warrants a closer look.

2.3. The special case of intelligence

2.3.1. The relevance of intelligence for OP

Intelligence can be defined as “a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience” (Gottfredson, 1997a, p. 13). Although there are differences in the definition of intelligence, especially with regard to its scope, a number of traits are seen as central to intelligence unanimously: Abstract thinking, problem solving ability, and the capacity to acquire knowledge seem to be central parts of intelligence (Snyderman & Rothman, 1988, p. 56). With regard to OP, Numerous reviews and meta-analyses have established the strong link between intelligence and a wide range of occupational performance indicators: The relevance of intelligence for the area of organizational psychology spans across the life-cycle of individuals, across and within occupations and includes occupational indicators from leadership to wages, prestige of occupations and supervisory ratings (e.g., Gottfredson, 1997b; Judge, Colbert, & Ilies, 2004; Lang, Kersting, Hülsheger, & Lang, 2010; Levine, Spector, Menon, Narayanan, & Cannon-Bowers, 1996; Ree & Carretta, 2002; Saigado et al., 2003; Schmidt & Hunter, 2004; Schmidt, Hunter, Outerbridge, & Goff, 1988).

---

[9] Prediction in the sense of statistical explanation of variance, not temporal precedence or causality.

Generally, analyses such as the prominent review of findings conducted by Schmidt and Hunter (2004) have shown the overarching relevance of intelligence for predicting successful human functioning ranging from primary education to old age (Gottfredson, 1997b; see also Schooler, Mulatu, & Oates, 1999, for a developmental perspective on the complex interplay between intellectual functioning and the work environment). More specifically for OP, Schmidt and Hunter present findings from several meta-analyses with corrected correlations between intelligence and performance on the job ranging from $r = .31$ to .73 and even stronger relations to training success ($r = .50$ to .76). The strong relation between intelligence and job performance even exceeds the importance of personality factors (see above) in terms of predicting occupational attainment and success (Schmidt & Hunter, 2004).

Even more, valid assessments of intelligence can be considered the single best predictors of occupational performance when selecting employees, going beyond selection methods originally developed for this purpose, such as Assessment Centers or structured interviews (Schmidt & Hunter, 1998). Importantly for this thesis, there seems to be a considerable conceptual overlap with CPS as indicated by the frequent use of problem solving and explicit references to complexity in definitions of intelligence (e.g., Gottfredson, 1997b; Raven, 2000; see also Oh, 2015).

According to a summarizing editorial by Linda Gottfredson signed by a range of central researchers of intelligence, the construct “reflects a broader and deeper capability for comprehending our surroundings – ‘catching on,’ ‘making sense’ of things, or ‘figuring out’ what to do.” (Gottfredson, 1997a, p. 13). This statement already points in the direction of the core processes of CPS, namely knowledge acquisition and knowledge application (Fischer et al., 2012, see Chapter 2.1). Remarkably, the importance of intelligence has been found to increase with rising complexity of the environment, culminating in the notion that “*g* is the ability to deal with complexity” (Gottfredson, 1997b, p. 93). Even more, reasoning as one of the most important factors of intelligence includes mental operations, such as drawing inferences, generating and testing hypotheses, identifying relations, comprehending implications, extrapolating, and transforming information (McGrew, 2009, p. 5). These operations of reasoning are corresponding to the main mental operations applied in CPS (Fischer et al., 2012; Greiff, Fischer, Stadler, & Wüstenberg, 2015, see Chapter 2.1, above).

A clarified relation between intelligence and CPS is therefore of utmost importance when evaluating the potential role of CPS in the organizational arena (Gottfredson, 1997b; Raven, 2000; Süß, 1996; see below). If the overlap of both constructs warrants no valid separation of CPS from (sub-facets of) intelligence, the area of organizational psychology will hardly benefit from insights generated in CPS research (but see Reeve, 2004, for a critical account of previous findings in that regard).[10]

And indeed, the research tradition of CPS has focused considerable energy on the disentanglement of CPS performance from established factors of intelligence (e.g., Dörner & Kreuzig, 1983; Stadler, Becker, Gödker, Leutner, & Greiff, 2015; Süß, 1996; Wüstenberg, Greiff, & Funke, 2012). Nevertheless, developments in the area of CPS assessment instruments (see also Chapter 3), as well as shortcomings of earlier studies have updated the need of investigating the constructs' relation. Hence, in view of the construct's central importance in the domain of organizational psychology, the first paper of this thesis is focused on the clarification of the relation between CPS and intelligence.

2.3.2. The relation of CPS and intelligence (Review Core Paper 1)

Kretzschmar, Neubert, Wüstenberg & Greiff (2016) provide an empirical examination of two competing theoretical perspectives on the relation of CPS and intelligence: The redundancy and distinctness perspective (see also Kretzschmar, 2015, for a more elaborate discussion sans the relation to OP). Starting from the overlap of mental operations characterizing both constructs, such as drawing inferences, generating and testing hypotheses, and identifying relations (see above), the perspective of redundancy conceptualizes CPS as a bundle of already available and established constructs (see e.g., Süß, 1996, 1999). In contrast, the distinctness perspective is focusing on the unique requirements of CPS. In the latter perspective, the overarching importance of intelligence and *g* for the handling of complexity is acknowledged,[11] but CPS is conceptualized as an umbrella term for all the requirements resulting from complex problem situations also including non-cognitive

---

[10] The issue of sub-facets of intelligence being weaker related to occupational performance and attainment and the potential consequences for CPS research in the domain of organizational psychology will be discussed below.

[11] Importantly, the perspective thereby differentiates from early accounts in CPS research, that argued for a complete independence of CPS from intelligence (e.g., Dörner & Kreuzig, 1983).

aspects, such as the need for emotion regulation and meta-strategic knowledge (e.g., Fischer et al., 2012; Fischer & Neubert, 2015; J. Funke, 2010).

Both perspectives provide empirical evidence in their support while also exhibiting differences related to CPS assessment instruments, operationalizations of intelligence, the consideration of external criteria, and the utilization of statistical methods of analysis. Consequently, Kretzschmar, Neubert et al. (2016) combined the advantages of previous empirical studies from both perspectives into a comprehensive examination of the construct's relation: (1) CPS was assessed with semantically abstract multiple complex systems, that is, with the help of MICRODYN and MICROFIN (see Chapter 3 and Fischer, 2015). (2) Intelligence was operationalized in a broad way via the Berlin Structure Intelligence test (BIS test, Jäger, Süß, & Beauducel, 1997), allowing for the measurement of intelligence in a faceted way (operation factors: reasoning, mental speed, memory, creativity, as well as a general factor *g*).[12] Additionally, a test of general knowledge as an indicator of crystallized intelligence was included in assessment (i.e., the BOWIT test of general knowledge, Hossiep & Schulte, 2008). (3) The relation to external criteria was tested via school grades (i.e., grade point average) as indicator of academic achievement to allow for an evaluation of predictive validity.[13] Finally, (4) analyses utilized two alternative measurement models building on structural equation modeling and the analysis of latent variables (i.e., a first order factor model and a nested-factor model).[14] With the help of this consolidated approach, Kretzschmar, Neubert et al.'s (2016) study clarifies the relation of CPS to established constructs of intelligence.

The empirical results of the study building on the participation of (N = 227) university students are indeed interesting: Generally and as could be expected from previous studies, there was a strong relation between intelligence and CPS, but the strength of association varied for the different operation factors: When predicting CPS from the established factors of intelligence in a latent regression model, reasoning (β = .85, p = .01) and creativity (β = -.34, p = .04) were significant predictors of CPS, while mental speed (β = -.02, p = .94), memory (β = .19, p = .16), and general knowledge (β = −.06, p = .56) were not. The negative relation between

---

[12] Differences in the three content factors of the BIS test (i.e., figural, numerical, verbal) were not analyzed in Kretzschmar, Neubert et al. (2016).
[13] Predictive in the sense of explanation of variance, see Footnote 9.
[14] Please see the full paper for more details on Kretzschmar, Neubert et al. (2016).

creativity and CPS in the regression model together with a low zero-order correlation points towards a suppressor effect of creativity. Overall, 60.2% of the variance in CPS was explained by the factors of intelligence.

The existence of a dedicated CPS factor was further investigated by the analyses of the nested factor model, where an explicit modeling of *g* via the shared variance of a reference construct (i.e., reasoning) and the specific abilities is combined with specific factors indicating the unique systematic variance of each construct. Results indicated substantial factor loadings on the *g*-factor for CPS ($Mdn(\lambda) = .59$), comparable to those of the other specific factors ($Mdn(\lambda)$ ranging from .53 to .68) and a specific reliability ω_s of .41, below the recommended benchmark (other specific factors $\omega_s = .41$ to .58). Importantly, a model without the specific CPS factor showed significantly worse model fit: $\Delta\chi^2 = 27.866$, $df = 1$, $p = .01$; ΔBIC = 17.

The results for the prediction of school grades were surprising, as CPS did *not* explain additional variance in school grades beyond the established factors of intelligence. More specifically, reasoning ($\beta = .57$, $p = .01$), mental speed ($\beta = -.58$, $p = .01$; as suppressor), and memory ($\beta = .39$, $p = .01$) were significant predictors of school grades, but creativity ($\beta = .20$, $p = .05$), general knowledge ($\beta = -.05$, $p = .32$), and CPS ($\beta = -.04$, $p = .41$) were not. Compared to a model without CPS, the amount of additionally explained variance in school grades was not significant ($R^2 = 32.6\%$, $\Delta R^2 = 0.7\%$, $F(1220) = 2.28$, $p = .13$). In contrast, adding CPS to the prediction of school grades by a narrow operationalization of intelligence (i.e., figural reasoning) led to a significantly better prediction of the external criterion ($R^2 = 10.1\%$, $\Delta R^2 = 3.9\%$, $F(1224) = 9.78$, $p < .01$). Similar results were found for the nested factor model, where only a narrow operationalization of intelligence via figural reasoning led to the significant prediction of school grades by the specific factor of CPS ($\beta = .18$, $p = .04$, $R^2_{total} = 10.1\%$), but not in the case of the broad model ($\beta = -.02$, $p = .43$, $R^2_{total} = 34.4\%$).

The results of Kretzschmar, Neubert et al. (2016) are an important clarification of the relation between CPS and intelligence. The results with regard to established factors of intelligence, such as reasoning and mental speed point to the separable, but strongly related nature of CPS: Similar to previous empirical research from the distinctness perspective utilizing figural reasoning tests (e.g., Greiff, Fischer, et al., 2013; Neubert, Kretzschmar, et al., 2015; Wüstenberg et al., 2012), a separable factor of CPS could be also established in the nested-factor model when utilizing a construct-valid operationalization of intelligence. Nonetheless, the high amount of explained variance in CPS (i.e., R^2 = 60.2%), factor loadings on the general factor *g* comparable to the ones of established factors of intelligence, and a latent correlation with reasoning of r = .72 point to the strong relation of CPS to intelligence. In fact, similar findings were interpreted as an indication of convergent (Kröner, Plass, & Leutner, 2005; Süß, 1996) or discriminant validity (Wüstenberg et al., 2012) in previous studies on the relation of CPS and intelligence and match the relation usually found for different factors of intelligence (e.g., Heller, Kratzmeier, & Lengfelder, 1998). Hence, it remains unclear, whether to position CPS as a (sub-)factor of intelligence or something beyond (e.g., in terms of a competency, see Fischer & Neubert, 2015, see Chapter 5).

What seems clear from both, the conceptual overlap of intelligence and CPS and also the empirical results by Kretzschmar, Neubert et al. (2016) is the overarching importance of intelligence for the handling of complex problems. This prominent role of intelligence might also limit the influence of other constructs when looking more broadly at CPS, for example via competency models (Fischer & Neubert, 2015; J. Funke et al., in preparation; see also Chapter 5).

With regard to the domain of OP, the findings concerning the interrelation of CPS and intelligence also have important implications. The strong relation between CPS and intelligence is especially noteworthy in the context of the central role of intelligence for OP (see above). On the one hand, the overlap of CPS and intelligence and the empirical dominance of intelligence in the domain of OP point towards the overarching importance of handling of (complex) problems in the domain of work (see Chapter 4). The finding of a separable CPS factor that can be differentiated from established factors of intelligence promises additional explanations detailing the affordances and processes of human interaction with complex problems in the domain of OP in that regard.

On the other hand, the results also point to potential problems regarding research on the role and implications of the specific aspects of handling complex problems captured by CPS instruments. In the study by Kretzschmar, Neubert et al. (2016), CPS did not show additional explanatory value in predicting school grades as markers of academic achievement.[15] Future studies building on construct-valid operationalizations of intelligence in the domain of OP will have to show, whether CPS can achieve explanatory value in addition to established factors of intelligence in this context. Similar to the relevance of a construct-valid operationalization of intelligence for the study by Kretzschmar, Neubert et al. (2016), these investigations will need to be built on theoretically and empirically trustworthy instruments indicating CPS performance. The following Chapter is therefore taking a closer look at the measurement instruments targeting CPS.

[15] Importantly, additional variance in school grades could be explained when comparing CPS to a narrow operationalization of intelligence as usually found in previous studies from the domain of CPS research (e.g., Wüstenberg, Greiff, & Funke, 2012), pointing to problems on the level of operationalizing intelligence adequately in these studies.

3. The assessment of CPS

3.1. Historical overview on CPS assessment

3.1.1. Classical studies and instruments

The founding fathers and mothers of CPS research utilized the availability of computers to simulate complex problems in a range of settings, thereby "escaping the narrow straits of the laboratory and the deep blue sea of the field study" (Brehmer & Dörner, 1993, p. 171). The prominent example of the TAILORSHOP simulation was already described in Chapter 1, simulating the management of a shirt factory for the purpose of investigating human interaction with complex problems.

The instruments utilized in early CPS research include a range of different computer simulations such as the already mentioned simulations of small towns, factories, and developmental aid (see Chapter 1, see also e.g., J. Funke, 1992; Gonzalez et al., 2005; Wallach, 1998, for overviews and taxonomies). Importantly, these simulations already feature some close links to practice in applied (industrial) settings and thus OP: In POWERPLANT, problem solvers have to control a coal-fired power plant modeled after the example of an existing one in Germany and hence, an actual work place from the domain of energy production. Even more, the processes of the simulation were chosen with the explicit goal of matching the requirements of operators in the real power plant (Wallach, 1998; Wallach & Tack, 1998). Other examples include the use of modified training simulations developed for the education of air traffic controllers (Ackerman, 1992), but also the simulation of everyday technical appliances, such as video recording equipment (Gray, 2000; see also Beckmann & Goode, 2013; and Neubert, Lans, Mustafic, Greiff, & Ederer, 2017, for more examples). While these investigations promise interesting insights for the connection of basic cognitive processes relevant for CPS and OP by fleshing out the different components of a CPS competency or "systems competency" (Fischer & Neubert, 2015; J. Funke et al., in preparation; see also Chapter 1), the focus in the following is on a second development here, the development and validation of instruments directed at the assessment of problem solving competencies.

3.1.2. Formal models and multiple complex systems

The investigation of CPS as a psychological construct as well as its application in OP is not only depending on sound theory and an empirical investigation of its nomological network, but also on the availability of reliable and valid assessment instruments. Building on the modeling of complex problems via computer simulations, major developmental milestones in this respect were the introduction of formal models to the investigation of CPS processes (e.g., J. Funke, 2001) as well as the use of multiple and smaller complex systems for assessment (Fischer, 2015; Greiff, Fischer, et al., 2015; Greiff et al., 2012).

With the help of formal descriptions of the (mathematical) structure underlying simulations such as the TAILORSHOP, different problems can be compared to each other on the basis of explicit criteria, thereby overcoming the problem of whether one should “attribute experimental findings to the experimenter’s manipulation or to the peculiarities of the task employed” (J. Funke, 2001, p. 70, see also J. Funke, 1992; Kluge, 2008a). To overcome this problem and provide an objective basis for performance evaluation, Joachim Funke introduced the formal frameworks of linear structural equations (LSE) and finite state automata (FSA) to the assessment of CPS competencies (J. Funke, 2001; see also Buchner & Funke, 1993; J. Funke, 1985). With the help of these formal frameworks, assessment instruments targeting problem solving competencies could be systematically developed with regard to different factors, such as varying difficulty by system size or the inclusion of different functional components of complex problems (e.g. including different problem features, such as osscilatory eigendynamics; cf. Hundertmark, Holt, Fischer, Said, & Fischer, 2015) as well as scored objectively with regard to successful problem solving (but see e.g., Dörner, 1989b, for a discussion of the inherent problem of scoring performance on complex tasks).

Furthermore, the use of formal models also allowed for the comparison of simulations including different content labels, while controlling for the systems structure. To this end, Beckmann and Goode (2013) investigated the (potentially) negative influence of semantic embedding and perceived familiarity of labels on problem exploration via untested assumptions. Through the use of content lean problem embedding (e.g., the use of labels such as ‘Variable 1’), the problem of differences in knowledge relevant for exploring and controlling complex problems could also be overcome (e.g., Kluge, 2008a; Wirth & Funke, 2005).[16] Overall, the availability of formal models facilitated the assessment of problem solving

[16] An alternative approach to the controlling of knowledge is the utilization of static knowledge tests during assessment, such as the one proposed by Süß (1996) for the Tailorshop simulation.

competencies by providing a systematic basis for formal description and development of instruments.

A second important development towards viable assessment instruments targeting problem solving competency was the introduction of multiple smaller simulations within one assessment session (Multiple Complex Systems, MCS; see also Fischer, 2015; Greiff, 2012; Greiff, Fischer, et al., 2015; Greiff & Funke, 2009; Greiff et al., 2012).[17] Larger simulations, such as the TAILORSHOP, typically take more than 30 minutes to complete, while mostly providing single criteria for evaluation (e.g., total capital of the company at the end of the simulation, but see Engelhart, Funke, & Sager, 2013). Researchers and practitioners interested in CPS were therefore forced to choose between assessment sessions lasting up to several hours for CPS alone or the use of single behavioral criteria (so-called single act criteria, see Fishbein & Ajzen, 1974).[18] To overcome this suboptimal choice, the instruments comprised under the label of MCS combine up to 10 smaller simulations each requiring 5-6 minutes into one session of assessment (e.g., instruments such as MICRODYN, Greiff & Funke, 2009; GENETICS LAB, Sonnleitner et al., 2012; and MICROFIN, Greiff, Fischer, et al., 2013; Neubert, Kretzschmar, et al., 2015).

Building on the two formal frameworks introduced by Funke (2001), namely LSE and FSA, the combination of multiple smaller complex problems thereby promises several advantages for CPS competency assessment (see Greiff, 2012; Greiff, Fischer, et al., 2015): The combination of different smaller simulations of varying size and complexity allows for the scaling of instruments with regard to difficulty, providing reliable estimations of CPS competency for individuals differing along the competency range. Furthermore, the psychometric requirement of stochastically independent indicators can be easily met, preventing problems in parameter estimation and providing straightforward estimates of reliability through internal consistency. Finally, the impact of (random) errors at the beginning of assessment can be reduced (Fischer, 2015; Greiff, 2012; Kretzschmar, 2015; see J. Funke, 2014a, for a more critical view on the MCS approach).

---

[17] The approach has been also referred to as 'minimal complex systems' in the beginning (Greiff, 2012; Greiff & Funke, 2010), a name that points to the potential problem of different requirements compared to more complex problems (J. Funke, 2014a; Kretzschmar, 2015, see also Chapter 5.2).

[18] The alternative approach of combining several (independent) indicators within one single simulation (e.g., Wagener, 2001) overcomes some of the problems mentioned here but still suffers from psychometric problems (Greiff, 2012; Greiff, Fischer, Stadler, & Wüstenberg, 2015).

3.1.3. Assessment instruments building on MCS

The two major proponents of MCS instruments available at the beginning of this dissertation project, GENETICS LAB and MICRODYN, are building on a combination of the formal framework of LSE and the MCS approach (Greiff & Funke, 2009; Sonnleitner et al., 2012). Both instruments build on the popular DYNAMIS framework and participants have to explore the (mainly linear) relations between two to three input and output variables, find out about relations between the variables and use their acquired knowledge to bring the systems' output variables into a given output range. Within a first phase, participants are given the opportunity to freely explore the task, before being asked to draw relations between variables into an abstract causal model (see J. Funke, 1992, for more details on knowledge assessment in CPS). In the final phase of each task, participants are given the correct structure of the task and asked to reach predefined values in the output variables, thereby targeting the application of knowledge. The phases of assessment are thereby structured along the assumed processes of CPS (see Chapter 2 and Fischer et al., 2012), and scoring of performance differentiates exploration behavior, knowledge acquisition and knowledge application (see Kretzschmar et al., 2016; Neubert, Kretzschmar, et al., 2015, for details on the MICRODYN problems utilized in the empirical studies of this thesis, including the underlying structural equations). The underlying problem relations in LSE-based MCS instruments are thereby defined via a set of mathematical equations relating each output variable to its previous states, as well as those of the input variables.

With regard to psychometric features, the instruments building on LSE and MCS have shown very promising results with regard to theoretically aligned measurement models, internal consistency, and measurement invariance across groups (e.g., Greiff & Wüstenberg, 2015; Greiff et al., 2012). Building on this basis of psychometrically solid features, the availability of assessment-oriented instruments targeting CPS has provided the means for a range of insights with regard to the construct of CPS, for example clarifying the constructs' relation to other established psychological constructs, such as intelligence and personality (e.g., Greiff & Neubert, 2014; Kretzschmar et al., 2016; Wüstenberg et al., 2012, see Chapter 2), but also constructs with a stronger link to application, such as school competencies, ICT literacy, and entrepreneurial opportunity identification (e.g., Baggen et al., 2015; Greiff, Kretzschmar, Müller, Spinath, & Martin, 2014; Kretzschmar, Neubert, & Greiff, 2014, see also Chapter 4). Additionally, the shorter testing times and psychometric properties have also allowed for the inclusion of CPS in international large scale assessments in the educational domain, such as the Progamme for the International Student Assessment (PISA, OECD, 2014).

Nevertheless, there are also challenges associated with the available instruments. One of these challenges is the reliance on a very similar set of tasks and hence, similar requirements in terms of exploration, knowledge acquisition and control in instruments building on the combination of LSE and MCS. Both MICRODYN and GENETICS LAB focus on a very specific type of complex problem in terms of identifying mainly linear relations between up to four input and output variables. And although the combination of MCS and LSE theoretically allows for the description and simulation of problems with a far greater range of problem characteristics, typical applications only feature a very limited number of linear relations and linear eigendynamics (see for example the equations used in Greiff & Neubert, 2014; Greiff et al., 2012; Kretzschmar et al., 2016; Wüstenberg et al., 2012; but see Hundertmark, et al., 2015, for a very interesting extension to "oscillatory eigendynamics" and random effects).

On the level of required strategies, this homogeneity in simulated tasks or more specifically, their underlying structure, leads to an overwhelming importance of one specific input sequence, a modified variant of the vary-one-thing-at-a-time (VOTAT) strategy (Greiff et al., 2012; see Chen & Klahr, 1999; Tschirgi, 1980; Vollmeyer, Burns, & Holyoak, 1996, for details on VOTAT). Utilizing the VOTAT strategy requires participants to systematically test the effect of isolated inputs on output variables to identify the system's relations (see Rollett, 2008, for an encompassing analysis of a DYNAMIS-based task).[19] For example, a participant first observes the simulation without manipulations for one round, in the next round increases the level of the first input variable and observes the effects on the output variables, then resets the simulation and manipulates the second input variable and so on.

[19] Importantly, scoring exploration in MICRODYN tasks is limited to single input steps, with "strategies" resulting from the scoring of specific input steps without considering sequence. Rollett (2008) introduced a more comprehensive model of scoring to DYNAMIS-based tasks, but his model has yet to be applied in the context of MCS-based tasks.

Empirically, the dominance of the VOTAT strategy in addition to rounds without manipulations for the successful exploration of most CPS instruments has led to results speaking against a separation of applying this specific exploration strategy (i.e., VOTAT application) and successful knowledge acquisition in general in most studies (e.g., Greiff & Neubert, 2014; Wüstenberg et al., 2012; see Kretzschmar et al., 2016; Kretzschmar, 2015, for a discussion of measurement models). The reliance on instruments basically 'only' requiring the application of VOTAT for exploration has therefore led to the criticism of CPS instruments building on MCS only representing an important but nonetheless small and non-representative sample of complex problems and thus, a non-representative assessment of CPS competency (Fischer, 2015; J. Funke, 2014a; Kretzschmar, 2015; see Fischer & Neubert, 2015; J. Funke et al., in preparation, for an interpretation in the context of more encompassing competency models). Especially important in this regard is the need for multimethod studies for a sound evaluation of construct validity (e.g., Oh, 2015).

And while the VOTAT strategy has a close resemblance to successful hypothesis-driven exploration in scientific settings (e.g., Chen & Klahr, 1999; Scherer & Tiemann, 2012), successful handling of complex problems conceptually requires more 'complex cognition' than the application of a single albeit important exploration strategy (J. Funke, 2010; see also the importance of adapting knowledge acquisition strategies already present in the work of Jean Piaget, e.g., Kuhn et al., 1995). Additionally, research on causal reasoning has already pointed towards the difficulty of controlled comparisons for knowledge acquisition in complex problem situations, such as those typically found in the context of OP (see Chapter 2.1 and Kuhn et al., 1995), which underlines conceptual problems when relying on VOTAT alone during CPS assessment.

In the history of problem solving research, the failed attempt at a general problem solver sensu Ernst and Newell (1969) building on a single strategy (i.e., means-ends-analysis) are speaking for the need to adapt problem solving attempts to the specifics of a situation (e.g., McDermott, 1976; Ohlsson, 2012), and hence, the need to account for different affordances of problems in assessment (see also the differentiation by Rollett, 2008, separating problem specific and problem general strategies). Similarly, practice in all kinds of applied settings including OP requires a much broader set of exploration strategies, heuristics, and problem solving efforts adapted to the current problem situation (see e.g., Jonassen, 2004, and the broad range of problems discussed there). For example, everyday problem solving of managers rarely follows scientific procedures of problem solving (see Mintzberg, 1973, 1975; see Yukl, 2010, for a more recent overview) and difficulties in typical everyday complex problems already arise during the identification of (important) variables and the setting of a specific frame (see e.g., Schön, 1993). Finally, earlier simulations of complex problems, such as the TAILORSHOP are hardly controllable with the VOTAT strategy alone, leading to the notion of MCS instruments and more complex simulations assessing different constructs altogether (J. Funke, 2014a). In summary, the reliance on a small selection of problems in assessment poses serious conceptual threats to a valid assessment of CPS and hence, its connection to OP.

To overcome this homogeneity in assessment instruments building on MCS and LSE, the second core paper in this dissertation therefore aims at expanding the range of available assessment instruments targeting CPS, while utilizing the benefits of the MCS approach. To this end, the second formal framework proposed for the description of CPS assessment instruments by Funke (2001), finite state automata (FSA), is combined with the MCS approach to construct a more heterogeneous assessment of CPS (Neubert, Kretzschmar, et al., 2015; see also Kretzschmar, 2015).

3.2. Assessing CPS competencies with MCS and finite state automata (Review Core Paper 2)

Neubert, Kretzschmar, Wüstenberg, and Greiff (2015) build on the notion of MICROFIN[20] introduced by Greiff and Funke (2009) as a combination of the MCS approach with the formal framework of FSA and present an assessment instrument targeting CPS (see also Fischer, 2015; Greiff, Fischer, et al., 2013; Kretzschmar, 2015). Contrasting to the established instruments building on MCS and LSE (see above), the use of a FSA as the formal framework facilitates the simulation of new problem features within the established framework of MCS, for example including strong interactions of input variables (such as those found in chemical problems,

[20] The name MICROFIN is thereby derived from the framework of FINite state automata (see Buchner & Funke, 1993), similar to MICRODYN building on the DYNAMIS framework.

Scherer & Tiemann, 2012) or qualitative changes in the problems' reaction to input changes after threshold values (e.g., water freezing to ice).[21]

Generally speaking, modeling problem solving via FSA views them as exhibiting a set of predefined states that are connected to each other via transitions. Transitions between states are either triggered by input signals or autonomous processes (e.g., passage of time) which in turn lead to output signals depending on the state the finite automaton was in. For example, a lighting system exhibits two states (light on/off), transitions between both states are triggered by pressing the light switch and the states are directly connected to corresponding outputs (light/no light). Larger finite-state automatons may include a broader set of input variables (e.g., setting a timer on the light), a broader set of states with corresponding transition rules (e.g., different light colors or brightness), and states or state-transitions not directly visible through changes in outputs (e.g., turning off an automatic cycle after manual activation of the light). Outside of CPS research, finite-state automata are widely applied in the programming of appliances (e.g., vending machines or turnstiles, J. A. Anderson, 2006; Rich, 2008). These applications of FSA have also led to visualization techniques via state-transition-diagrams (e.g., Figure 2 in J. Funke, 2001 for the formal description of a very small finite state automaton; or Appendix A1 in Neubert, Kretzschmar, et al., 2015, the state-transition-diagram of an exemplary MICROFIN task).

With regard to CPS research, the formal framework of FSA was already introduced by Buchner and Funke in 1992, who also present ways to model FSAs during assessment and proposed several approaches to knowledge measurement (see Buchner & Funke, 1993; J. Funke & Buchner, 1992). Initial applications of FSA provided valuable foundations for CPS research, for example in the national extension of the PISA studies in Germany (e.g., Klieme, Funke, Leutner, Reimann, & Wirth, 2001; Leutner, Fleischer, Wirth, Greiff, & Funke, 2012; Leutner, Klieme, Meyer, & Wirth, 2004). Nonetheless, combinations of FSA and MCS with its advantages described above were either restricted to theoretical explorations (e.g., Greiff &

[21] Please note that the formal framework of LSE in principle allows for the simulation of these and other phenomena in complex problems, but assessment instruments did only include a narrow set of problem features (see above) and the inclusion of features, such as threshold values or tipping points would require adaptations to assessment (e.g., during knowledge assessment). See for example Hundertmark et al. (2015) for a recent extension of the LSE approach to non-linear dynamics in assessment.

Funke, 2009) or suffered from a very restricted sample of tasks (e.g., Greiff, Fischer, et al., 2013, only included two tasks in assessment).

To overcome this lack of instruments building on the combination of FSA and MCS, Neubert et al. (2015) developed, piloted, and refined a set of MICROFIN tasks (see also Kretzschmar, 2015).[22] The explicit goal of test development was the extension of requirements beyond the focus on VOTAT as found in available LSE-based instruments while retaining the advantages of the MCS approach. That is, the tasks were targeted for an application within the MCS framework, so the size of problems, as well as their reliance on prior knowledge (e.g., concerning business processes) was restricted: Testing times for all tasks are similar to those found in established LSE-based tests with around 5 minutes per task and the influence of prior knowledge is restricted through knowledge lean descriptors. The extension of exploration strategies is best presented with an example.

An exemplary MICROFIN task, called “Fish-o-maton” can be seen in Figure 1 (see also Neubert, Kretzschmar, et al., 2015). In the Fish-o-maton, participants have to explore the effects of varying the input levels in three tanks seen at the bottom of Figure 1 (i.e., three ordinal input variables) on an aquarium (i.e., one nominal output variable with ordinal elements). The underlying scheme of relating input to output variables thereby follows the example of LSE-based instruments, such as MICRODYN or TAILORSHOP. The important extension is the type of relation: Fish are only seen in the aquarium if all input variables are put on equal levels (as seen in Figure 1), otherwise the aquarium is in a “soiled” state or completely empty (if all input variables are put to the lowest possible value). The input variables each have four possible input values (tanks being empty or filled in three stages), while the output variable has five possible values (empty, soiled, and three states with an increasing number of fish in the tank). Due to the relation of input and output variables, the application of the VOTAT strategy leads to suboptimal results in the Fish-o-maton: If all input variables are varied in isolation from zero to three, the central states of all inputs being on the same level is omitted. Hence, the task extends the range of necessary exploration behaviors.

[22] The MICROFIN tasks have been developed in collaboration by André Kretzschmar and Jonas Neubert, the first two authors of Neubert, Kretzschmar et al. (2015).

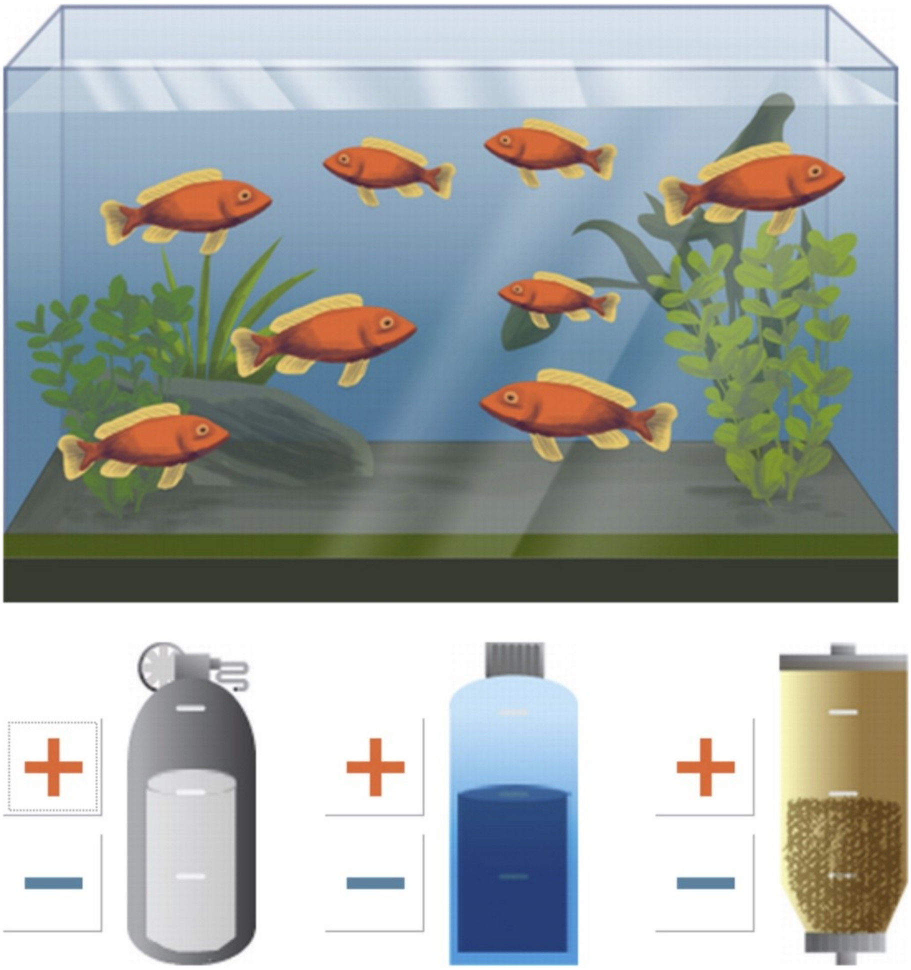

Figure 1. Screenshot of the MICROFIN task Fish-o-maton (see Neubert, Kretzschmar, et al., 2015, p.183).

The set of tasks developed by Neubert, Kretzschmar et al. (2015) extends the relations between input and output values as found in MICRODYN and GENETICS LAB with features, such as equivalence of inputs (see above), threshold values (i.e., the problem reacting differently to input variations if a certain value in inputs is reached), and subsystems following different rules. Further details on all MICROFIN tasks and formal descriptions can be found in Neubert, Kretzschmar et al. (2015) and an extensive description of the development process of MICROFIN is available in Kretzschmar (2015).

With regard to the construct of CPS, the newly developed MICROFIN tasks target the two dimensions of CPS already included in established LSE-based instruments, namely knowledge acquisition and knowledge application (see Chapter 2.1). That is, after freely exploring the tasks during an exploration phase for a maximum of 300 seconds, participants are asked to answer two questions on rules governing the system (knowledge acquisition) and bring the system into two different predefined states (knowledge application).

The questions utilized during knowledge acquisition assessment build on an approach introduced by Buchner and Funke (1993) and ask participants to construct the initial state of the system prior to a given change in inputs. For example, in the Fish-o-maton participants are given an output state of the aquarium featuring equal levels of inputs and fish in the aquarium and have to select the appropriate state of the aquarium prior to a user manipulation of one input variable (e.g., reducing the input level from maximum to equality with the other inputs). The correct solution has to be constructed from one to eight different elements, thereby assessing the rule knowledge acquired throughout exploration (i.e., a constructed response item, Buchner, 1995), in the given example, the correct solution would be the selection of a soiled aquarium. Scoring of knowledge acquisition in Neubert, Kretzschmar et al. (2015) is done by comparing the participants answer to the correct solution resulting in a sum score over both items of each task ranging from 0 to 2. During the development of MICROFIN, alternative forms of assessment were tested (e.g., the identification of possible states), relating to other forms of knowledge, but these alternatives have not been systematically evaluated so far (see also Kretzschmar, 2015).

In knowledge application, participants have to bring the respective task into a given target state within a time limit of 60 seconds by manipulating the input variables (i.e., triggering state-transitions) that is presented visually and verbally at the beginning of the phase. In the case of the Fish-o-maton, a knowledge application task might require participants to reach a state of eight fish in the aquarium with as few steps as possible. Again, two items are presented per task and participants receive credit for reaching the target state, resulting in a sum score ranging from 0 to 2 for knowledge application. In contrast to MICRODYN, no correct model is given before knowledge application in MICROFIN.

Scoring of participants' exploration behavior as a third dimension of CPS (i.e., the equivalent of scoring exploration via VOTAT in MICRODYN or GENETICS LAB) is much more difficult in MICROFIN due to the broader range of required exploration strategies (see above). For an initial attempt at scoring the exploration phase of participants in MICROFIN, see Müller, Kretzschmar, & Greiff (2013), who adapted the idea of VOTAT to the MICROFIN tasks in a strategy termed 'nested VOTAT' depending on an adaptation of exploration behavior to the respective task. Low reliabilities and conceptual problems in scoring strategies in the different MICROFIN tasks point towards the need for further development in this area for an explicit inclusion of exploration behavior in MICROFIN assessment. Assessment of MICROFIN has therefore been restricted to the assessment of acquired knowledge and the successful control of the systems.[23]

Empirically, Neubert, Kretzschmar et al. (2015) targeted (1) the psychometric properties of the newly developed set of MICROFIN tasks, (2) established a measurement model for MICROFIN, (3) analyzed the relation to MICRODYN as an established instrument targeting CPS with the help of the MCS approach (i.e., convergent validity), and (4) explored the relation to reasoning as a competing construct from the domain of intelligence (see also Chapter 2.3). The empirical analyses can build on data provided by 576 German high school students that worked on the newly developed MICROFIN tasks, MICRODYN tasks building on LSE and MCS, and an assessment of reasoning ability (i.e., the CogAT, a figural matrices test, Heller & Perleth, 2000).

With regard to the psychometric properties of MICROFIN, the results of Neubert, Kretzschmar et al. (2015) support the viability of combining FSA with MCS for the purpose of CPS assessment. Item difficulty as indicated by the average rate of success ranged from $p = .18$ to $.49$ ($M = .34$, $SD = 0.14$) for knowledge acquisition and from $p = .63$ to $.77$ ($M = .70$, $SD = 0.06$) for knowledge application. That is, the items targeting knowledge application proved to be easier than the ones targeting knowledge acquisition, pointing towards areas in need of further refinement. For the purpose of this dissertation, the sample of high school students and their prerequisites have to be kept in mind when inferring the difficulty of tasks for adult samples and applications in OP, although the students represent the full range of available school tracks in Germany. The reliability estimates of the five tasks point towards the general applicability of FSA-based tasks, with McDonald's omega of $\omega = .79$ (knowledge acquisition) and $\omega = .78$ for knowledge application, while also leaving room for improvement.[24]

---

[23] In applications of MICRODYN, too, an explicit inclusion of exploration behavior in analyses is mainly omitted, because of strong intercorrelations with knowledge application (e.g., Wüstenberg et al., 2012, see also above).

[24] When interpreting the values of McDonald's Omega (Zinbarg, Revelle, Yovel, & Li, 2005), one has to keep in mind the comparably low number of five tasks as well as the underlying dilemma of

The next step in exploring the utility of FSA-based MCS tasks is the establishment of a measurement model. Neubert, Kretzschmar et al. (2015) therefore compared the model fit of models featuring separate dimensions for knowledge acquisition and knowledge application, as well as a uni-dimensional model with a general CPS factor. Similar to the established models for MICRODYN, the two-dimensional model indicated a good model fit ($\chi^2(34) = 52.684$, $p = .021$, RMSEA = 0.031, CFI = 0.989, TLI = 0.986) with strongly related, but separable latent dimensions of CPS (latent $r = .81$, $p < .001$). Conflating the two dimensions of MICROFIN led to significantly worse model fit (χ^2 $(35) = 90.118$, $p < .001$, RMSEA = 0.053, CFI = 0.968, TLI = 0.958, $\chi^2\Delta(1) = 24.398$, $p < .001$). In terms of measurement models, the results of Neubert, Kretzschmar et al. (2015) point towards a two-dimensional measurement model for MICROFIN, highlighting the link to established instruments targeting CPS from the MCS perspective (but see Kretzschmar, 2015, for findings in favor of a unidimensional model).

The link to these established instruments was further examined empirically via a correlated trait-correlated method minus one model (CT-C(M-1); Eid, Lischetzke, Nussbeck, & Trierweiler, 2003, see Figure 2), a multitrait-multimethod model including the same latent dimensions of CPS for both instruments and a method factor for MICROFIN (i.e., the established instrument MICRODYN was used as the reference method). The model fit the data well (χ^2 $(285) = 445.247$, $p < .001$, RMSEA = 0.031, CFI = 0.980, TLI = 0.978). Again, the latent dimensions of CPS were strongly related ($r = .82$, $p < .001$), as were the latent method factors of MICROFIN ($r = .54$, $p < .001$). For a graphical impression of the model please also consider Figure 2. Overall, the model points towards the expected strong relation and dimensional structure of both instruments targeting CPS with its dimensions of knowledge acquisition and knowledge application. Similarly, the significantly related method factors for MICROFIN indicate a somewhat generalized method effect for the newly developed instrument differing from MICRODYN.

Interestingly, an analysis of a structural model featuring separate measurement models for both instruments shows the weakest relation between the instruments dimensions' targeting knowledge application ($r = .56$, $p < .001$), with

---

introducing more heterogeneous tasks to CPS assessment and the corresponding effect in terms of a reduced internal consistency across tasks (the “attenuation paradox”, Loevinger, 1954).

stronger relations of knowledge application to the other instruments' knowledge acquisition factor (*r* = .64 and *r* = .67, all *p* < .001) and between both factors indicating knowledge acquisition (*r* = .73, *p* < .001). Naturally, this second set of results raises questions with regard to the comparability of requirements in MICROFIN and MICRODYN.

To investigate the relation of both instruments to each other, as well as to reasoning ability, the multimethod model presented above was extended by a separate factor for the reasoning test (see Figure 2). The resulting model still fit the data well ($\chi^2(361) = 540.359$, $p < .001$, RMSEA = 0.029, CFI = 0.978, TLI = 0.976) and included significant relations between reasoning ability and both factors of CPS (*r* = .63, *p* < .001, for knowledge acquisition and *r* = .60, *p* < .001 for knowledge application), as well as between reasoning ability and the method factors for MICROFIN (knowledge acquisition, *r* = .34, *p* < .001, knowledge application, *r* = .28, *p* < .001). Interestingly, the relation between both dimensions of CPS did not decrease after including reasoning ability into the model in the CT-C(M-1) model, while the latent correlations in the model featuring separate measurement models for both instruments dropped to (*r* = .33 to *r* = .52, all *p* < .001, similar pattern as presented above), highlighting differences in results depending on the chosen measurement model. The results with regard to reasoning ability have to be interpreted with caution when generalizing to an effect of intelligence or *g* after the findings of Kretzschmar, Neubert et al. (2016) presented earlier (see Chapter 2.3), but they nonetheless point towards the viability of utilizing FSA-based instruments within the MCS framework to target CPS.

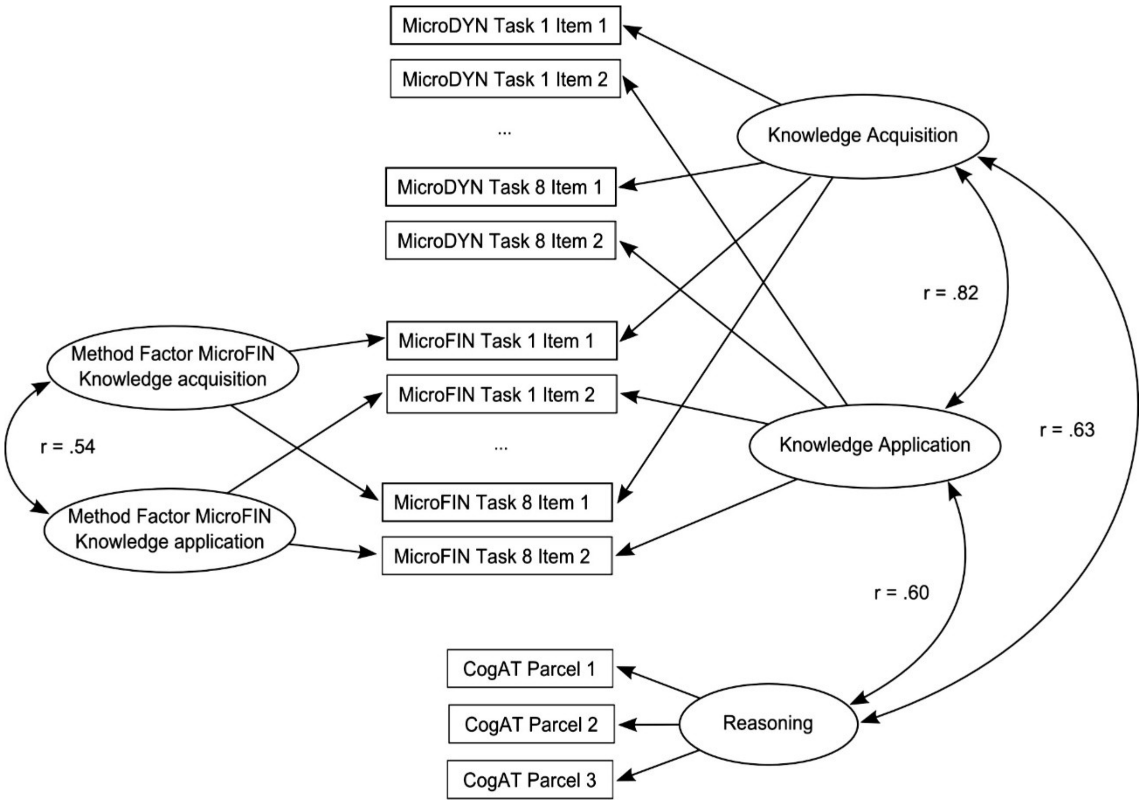

Figure 2. The CT-C(M-1) multitrait-multimethod model (see Neubert, Kretzschmar, et al., 2015, p.188). The model features latent factors for knowledge acquisition and knowledge application indicated by both MICRODYN and MICROFIN (right side), two method factors modeling specific aspects of MICROFIN per dimension (left side), and a latent factor indicating reasoning. Latent correlations for the method factors, error variances, and manifest indicators for some MICROFIN and MICRODYN tasks for better visual clarity.

The empirical results presented in Neubert, Kretzschmar et al. (2015) as well as the developed MICROFIN tasks are an extension of the MCS-based assessment of CPS. The availability of a broader set of MICROFIN tasks with promising psychometric features enables an extension of MCS-based CPS assessment beyond the strong focus on VOTAT. Since the publication of Neubert, Kretzschmar et al.'s (2015) article, more MICROFIN tasks have been developed and applied in a range of settings (for an overview see Kretzschmar, 2015). The separability of knowledge acquisition and knowledge application has been challenged in some cases (i.e., models with a uni-dimensional CPS construct fitting the data better, e.g., Kretzschmar et al., 2016), leading to the proposal of different measurement and structural models (Kretzschmar, 2015).[25] Similarly, the question of internal consistency has proven to

[25] The CT-C(M-1) model applied in Neubert, Kretzschmar et al. (2015) depends on a two-dimensional measurement model for both, MICROFIN and MICRODYN, so it is not applicable in cases where the optimal measurement model for MICROFIN is unidimensional or builds on an overall score on a task level.

be more challenging with levels as low as ω = .58 for knowledge acquisition and ω = .62 for knowledge application in an unpublished study with five MICROFIN tasks (see Kretzschmar, 2015). Indeed, the findings also point towards the (psychometric) advantage of using a restricted set of problem relations as found in MICRODYN in terms of higher reliability estimates. With regard to MICROFIN, the results show the need for further development, especially when looking for applications in the domain of individual assessment (e.g., within the setting of personnel selection requiring higher reliability and a broader difficulty range).

Conceptually, the availability of a broader set of tasks embedded within the MCS framework enables a broader sampling from the universe of potentially relevant problem features. One of the arguments for the development of MCS-based instruments was the possibility to construct assessment instruments targeted at specific and theoretically derived problem features and problem solving processes (e.g., Greiff et al., 2012). Unfortunately, the development of actual assessment instruments was strongly bound to a restricted sense of problem features and specifically a strong link to the application of the VOTAT exploration strategy (J. Funke, 2014a).

The developed set of MICROFIN tasks helps to overcome this focus, although not including a systematic sampling of problem features (see below). These extensions of the MCS framework to new problem features might lead to a better representation of requirements of complex problems as described by competency models of CPS (e.g., Fischer & Neubert, 2015; J. Funke et al., in preparation), but the MCS framework itself has been also criticized lately as a whole (e.g., J. Funke, 2014a). And indeed, in comparison to the multitude of approaches possible in more complex simulations, such as the FIRE simulations (De Obeso Orendain, 2014; Güss, Tuason, & Orduña, 2015) or other more comprehensive microworlds, the variability in behavior and strategies is severely restricted in MCS tasks. It remains to be seen, whether an extension of the MCS approach to other problem features can overcome these conceptual issues on the level of actual assessment instruments (see also Chapter 5.2).

From the perspective of OP, the availability of a range of problem features during the simulation of complex problems is mainly related to the question of establishing valid assessment instruments targeting CPS. The actual application of MICROFIN in the domain of OP and an empirical basis for the evaluation of its effect in terms of validity is still missing, as applications were restricted to samples of students in schools and universities so far (the application of MICROFIN in the LLLight'in'Europe project might offer insights in the future, see www.lllightineurope.com and Kretzschmar, 2015). On the one hand, the extension of MCS-based instruments to FSA might offer a potential link between practice in OP and a valid representation of complex problems in CPS assessment, especially given the widespread use of FSA in programming appliances. Sampling the tasks and requirements of complex problems in the domain of OP offers a route to the development of valid assessment instruments highlighting the various requirements of complex problems in OP, which in turn will rely on the availability to simulate a broad range of phenomena. The domain of OP, in turn, might benefit from the experience in describing and assessing human interaction with complex problems across various domains accumulated in CPS research (see Neubert et al., 2017).

On the other hand, analyses such as the ones conducted by Henry Mintzberg over 40 years ago point to the intricate nature of practically dealing with complex problems in the organizational domain. Problem solving there includes, for example, an emphasis on "soft" information, such as "gossip, hearsay and speculation" (e.g., Mintzberg, 1975, p. 14), which can be hardly considered to be a part in current CPS assessment. A broader view on CPS such as found in Funke et al. (in preparation) and competency models encompassing a range of requirements, such as proposed by Fischer and Neubert (2015) might offer a way to include these and similar features of complex problems and successfully handling them into CPS research, but it remains to be seen, whether they can be adequately described, simulated, assessed, and scored within the framework of MCS (see Chapter 5.2). At the moment, it is hard to imagine MICROFIN tasks adequately mirroring the nature and requirements of complex problems faced by managers in the domain of OP. Focusing on this interaction between current strands in CPS research and the domain of OP, Chapter 4 is therefore oriented towards an interdisciplinary exchange on the potential role of CPS in the domain of OP.

4. Complex Problem Solving and Organizational Psychology

The developments within the domain of CPS research have led to notable applications of assessment instruments and theoretical explorations in a different domain, namely that of education during recent years. As mentioned in Chapter 1, assessment instruments targeting CPS have found their way into educational large scale assessments, such as the Programme for the Internationel Student Assessment (PISA, OECD, 2014) and empirical studies are utilizing the construct with its nomological network and the corresponding instruments to explore the role and importance of CPS in educational settings (see Chapters 2 and 3). For example, Scherer and Tiemann (2012) highlight the interplay between domain general processes as explored in CPS research with domain-specific problem solving approaches in a virtual chemistry environment. Similarly, Kretzschmar et al. (2014) explore the role of CPS for successfully handling the requirements in school compared to subject-related competencies.

Strangely enough, the developments within CPS research, such as the interplay of intelligence and CPS or the MCS-approach have been met with much less response in the domain of OP. This lack of interdisciplinary exchange is somewhat surprising as the experimental findings from early CPS research, such as the typical problems of human problem solvers with complexity (e.g., a lack of identifying contradictory goals and insufficient model building, Dörner, 1989b) have had a definite influence on researchers investigating human problem solving and learning in the domain of OP (e.g., Fisch & Beck, 2004). Especially in the German OP literature, the results of Dietrich Dörner have been taken up by researchers focusing on a broad range of topics such as human errors and critical failures in organizations (e.g., Badke-Schaub, Hofinger, & Lauche, 2012; Strohschneider, 2007), knowledge management (e.g., Amelingmeyer, 2002; Probst, Raub, & Romhardt, 2006), consulting and organizational development (e.g., Schiersmann & Thiel, 2011), managerial and entrepreneurial decision making (e.g., Gustafsson, 2006; Wagner, 1991), systemic views on organizations (e.g., Schiepek & Strunk, 2006; von der Weth, 2001; Willke, 1992), and the design of decision making systems and procedures (e.g., Hacker & von der Weth, 2012; Sterman, 2006). In contrast, the development of psychometrically-oriented assessment instruments, the clarified importance of intelligence, or competency models describing human interaction with complex problems have seen no comparable transfer to the domain of OP (but see Abele et al., 2012; Rausch, Seifried, Wuttke, Kögler, & Brandt, 2016, for exceptions). The following Chapter therefore explores the potential of interdisciplinary collaboration between the domains of CPS and OP.

When taking a closer look at assessment instruments, the TAILORSHOP simulation as a representative of larger computer simulations used in CPS research has received some attention in assessment efforts in the organizational setting (e.g., Danner et al., 2011; U. Funke, 1993; Hasselmann, 1993; Kersting, 1999; Kluge, 2008a; Sonnenberg, 1993; see Süß, 1996, for an overview). For example, Danner et al. (2011) compared the performance in handling the TAILORSHOP simulation with an indicator of professional success (i.e., supervisor rating on a standardized questionnaire), highlighting the role of intelligence, but also the separable notion of a distinct problem solving competence (but see Kretzschmar et al., 2016). Following a different line of reasoning, there is a whole strand of research exploring the applicability of computer simulations for assessment in OP. For example, in 1995, Uwe Funke analyzed the utility of simulations such as TAILORSHOP for the purpose of management assessment, coming to rather mixed conclusions, due to an unclear link between the requirements of specific management jobs and the computer simulations used in CPS research (U. Funke, 1995; see also J. Funke, 1993; U. Funke, 1993; Hasselmann, 1993; Kersting, 1999; Sonnenberg, 1993; Wagener, 2001).

Nevertheless, interest in the latest developments within CPS research has not been corresponding to these earlier considerations, as well as the overcoming of some of the (psychometric) problems raised in earlier research looking at the applicability of computer simulations developed in the domain of CPS research within the domain of OP (e.g., Kersting, 1999). But why should the domain of OP care about the findings of CPS research to begin with?

4.1. Developments in the world of work

One of the reasons researchers and practitioners in OP might care about human interaction with complex problems and thus, new developments in CPS research, are the changes in the actual tasks humans perform at work (see e.g., Kersting, 1999, for a consideration of factors stronger related to the appeal to tested subjects). A number of authors from different domains have highlighted the intriguing development of human work activities during the last decades, as well as the accompanying shifts in requirements (e.g., Autor, Levy, & Murnane, 2003; Cascio, 1995; Howard, 1995; Spitz-Oener, 2006; Sterman, 2006). For example, economic researchers have highlighted trends towards non-routine and interactive work tasks contrasting to a decline in routine work (e.g., Autor et al., 2003; Autor & Price, 2013; Spitz-Oener, 2006).

That is, on the one hand, tasks defined by predefined procedures and repetitive action are declining in number and importance, for example during the preparation of standardized invoices for customers or the monitoring of production processes. On the other hand, tasks requiring the handling of non-routine and interactive situations are becoming more and more important, for example during the programming of the software preparing the invoices or when coordinating the needs of different organizational units involved in the monitoring of production in a restructuring project (see e.g., Middleton, 2002, for a description of this recursion to individual problem solving). Along the same lines, researchers within the domain of vocational education and tranining (VET), dedicated to the preparation of individuals for these tasks and jobs, have emphasized the change in requirements towards problem solving, innovation, and sustainability as one of the drivers of adopting competency frameworks in VET (e.g., Boreham, 2002; Brockmann, Clarke, & Winch, 2008; see also Neubert et al., 2017).

Different explanations for these changes have been raised, such as the computerization of routine tasks (e.g., Autor et al., 2003; Autor & Price, 2013), globalization and more specifically the offshoring of routine cognitive jobs (e.g., Baumgarten, Geishecker, & Görg, 2010; Becker, Ekholm, & Muendler, 2013; Grossman & Rossi-Hansberg, 2008), as well as changing customer demands (e.g., Autor & Dorn, 2013; see also Goos, Manning, & Salomons, 2009). Similar patterns can be observed throughout the western world, from the United States to Europe and Japan, highlighting the global importance of these developments (Autor, Katz, & Kearney, 2006; Goos & Manning, 2007; Goos et al., 2009; Ikenaga & Kambayashi, 2010; Spitz-Oener, 2006). Already on this level of changing work tasks, there is some overlap with the notion of CPS, namely the emphasis on handling problem situations that cannot be handled by simply recurring to predefined procedures (see Middleton, 2002), instead relying on the problem solving capabilities of employees and the need to consider ‘complexity’ (see Axley & McMahon, 2006; Tsoukas & Dooley, 2011;

Tsoukas & Hatch, 2001 for further explorations in organizational science; see also Chapter 5).

Within the domain of organizational research and OP, the changes in the world of work have been reflected for example in the literature on organizational routines looking at the rising variation due to computerization (e.g., Pentland & Hærem, 2015; Pentland, Hærem, & Hillison, 2011), the consideration of so-called 'wicked' problems and system dynamics in organizational planning and learning (Rittel & Webber, 1973; Sterman, 2001; see e.g., Duijnhoven & Neef, 2016, for a recent application in the area of operations research), and the extension of work and task analysis to non-routine and complex tasks and jobs (e.g., Jenkins, 2009; Naikar, Moylan, & Pearce, 2006; Vargas Cortes & Beruvides, 1996). Additionally, the awareness of the complex nature of problems might also explain the consideration of studies coming from the earlier days of CPS research in OP (see above).

Intriguingly, the U.S. Department of Labor's comprehensive overview on jobs and occupations, the Occupational Information Network (O*Net; N. G. Peterson & American Psychological Association, 1999, see www.onetcenter.org) also includes 'complex problem solving', defined as '[d]eveloped capacities used to solve novel, ill-defined prdoblems in complex, real-world settings' in their classification of worker requirements, highlighting the specific need to handle complex problems in some occupations . Although this inclusion signifies a general interest in the topic of dynamic work environments and the requirements they pose to individuals, it is nonetheless important to note that the O*NET classification is following a very different assessment and categorization of jobs and tasks compared to the performance assessment utilized in CPS research.

So in summary, there are examples of utilizing computer simulations, such as those employed in CPS research in the domain of OP and beyond, even if sometimes under different labels and for different purposes. And on the other hand there are developments in the world of work towards non-routine tasks, wicked problems, and the need to handle complex challenges also reflected in research and practice in OP with a strong overlap to CPS. The paper by Neubert, Mainert et al. (2015) therefore explored the potential of CPS research in OP more systematically and provided a starting point for a discussion with researchers and practitioners from the domain of OP by taking a look at the assessment of so-called 21st century skills, including CPS.

4.2. Exploring the potential role of CPS in OP (Review Core Paper 3)

A general argument for the consideration of complexity, complex problems, and how to deal with them in OP has been presented above, as well as previous efforts to include such notions in OP as well as insights from CPS. The question remains, whether there is a need to include recent findings from CPS research into the considerations in OP, given the alternatives available to the domain and the focus on practical interventions via its scientist-practitioner model (Bass, 1974; Dunnette, 1990; Murphy & Saal, 1990; Rupp & Beal, 2007). Neubert, Mainert, et al. (2015) therefore contrasted the approach taken in CPS research[26] to the assessment of CPS with three alternative ways to assessing prerequisites to handle the challenges described above: (1) Building up directly from the needs of practitioners and business leaders culminating in application-oriented constructs and instruments, (2) the use of systematic job and work analysis to specify the requirements of specific jobs or occupations, and (3) utilizing basic psychological constructs and the corresponding instruments, such as those targeting intelligence. Explorations of the opposite route of transfer, namely the integration of insights form OP into CPS research will be briefly explored in Chapter 5.

4.2.1. Exploring Alternatives to CPS Assessment

Generally speaking, the approach taken in application-oriented constructs departs from specific needs or problems with a high visibility and relevance for practitioners and organizations in the respective field and builds directly towards solutions for these needs or problems. Neubert, Mainert, et al. (2015) explore the benefits and problems of that approach by taking a closer look at the example of learning agility (De Meuse, Dai, & Hallenbeck, 2010; Eichinger & Lombardo, 2004; Lombardo & Eichinger, 2000).[27] In the example of learning agility, the point of departure is the need to identify individuals who will be able to successfully deal with the work environment characterized by dynamic changes and therefore the changing

[26] Neubert, Mainert, et al. (2015) also consider the case of Collaborative Problem Solving (ColPS) in their focal article, which will not be further examined here. This is also the case for the comment by Riggio and Saggi (2015) that is strongly focused on ColPS.

[27] A similar example would be the popular notion of talent management in current OP research (Cappelli, 2008; Ready & Conger, 2007; see Collings & Mellahi, 2009; Lewis & Heckman, 2006, for a critical view).

work environment depicted above (Eichinger & Lombardo, 2004; Lombardo & Eichinger, 2000; see also Neubert, Mainert, et al., 2015).

Proponents of the learning agility approach propose the construct as a means to identify high potential employees capable of coping with these situations, higher levels of learning agility indicating better chances of learning from experience in new or first-time conditions (De Meuse et al., 2010). Similarly, the strengthening of learning agility of employees is proposed as a means to cope with the increasing importance of non-routine tasks (Dries, Vantilborgh, & Pepermans, 2012). From this perspective, the application-oriented construct of learning agility seems to be a straightforward answer to the rising importance of non-routine tasks, directly rooted in a problem of high practical relevance and providing a specifically tailored answer.

Contrasting to this positive picture of learning agility as an example of application-oriented constructs, Neubert, Mainert et al. (2015) also highlight its' downsides: A lack of theoretical integration and severe assessment problems. With regard to theoretical integration and a fleshed out nomological network, learning agility's overlap with established constructs such as intelligence, personality, and cognitive styles remain unclear and empirically untested, leading to problems when targeting theoretical and empirical integration (cf. DeRue, Ashford, & Myers, 2012). For example, the unclear overlap with intelligence leads to problems when targeting higher order theories of how to cope with the changing work environment und thus, proliferates a fragmented view of phenomena only concerned with very specific problems and needs. Similarly, exploring the trainability of learning agility runs into problems when the unclear influence of intelligence and other constructs is not accounted for: A practitioner or organization will overestimate the effects of training learning agility on their employee's capabilities in dealing with complex problems due to an underestimation of the role of basic cognitive capabilities.

Problems related to the assessment of application-oriented constructs can be partially traced back to the missing theoretical and empirical integration and the resulting lack of well-tested, reliable, and valid instruments. For the example of learning agility, DeRue et al. (2012) identified serious problems when looking at a typical instrument targeting learning agility, such as conceptual confusion on the level of the instrument itself and basic errors in item construction such as double barreled items integrating two separate questions into one. In summary, Neubert, Mainert et al. (2015) conclude that compared to CPS assessment, dealing with the changes in work tasks via application-oriented constructs seems too strongly hampered by problems related to theoretical integration and valid assessment to offer a valid alternative.

The general problems associated with the separate 'worlds' of researchers and practitioners also visible in other examples of application-oriented constructs with little integration into research have been identified in OP as well, leading to ample discussions on ways to overcome the gap between both groups (e.g., D. J. Cohen, 2007; Rynes, Giluk, & Brown, 2007). For research on CPS, the example of application-oriented constructs highlights the possibility and benefits to start investigations into problem solving right at the needs and issues of practitioners, a possibility that has been also used in other domains of problem solving research, such as Naturalistic Decision Making (NDM, see G. Klein, 2008, for an overview). In the case of learning agility, the need of organizations to handle the changes in dynamic work environments in a strategic way becomes visible, as well as the need for insights on how to prepare individuals for these dynamic environments. In terms of CPS research, accounting for strategic importance of the topic in organizational settings seems beneficial and the presentation of the KSAO-model of CPS by Fischer and Neubert (2015) might offer a route of integration of this stronger link to practice in CPS research itself (see the application by Rausch et al., 2016).

The second alternative to deal with the rising complexity in work tasks discussed by Neubert, Mainert et al. (2015) is the utilization of work and job analysis already ingrained in the OP literature since Frederick Taylor (1911) and his Principles of Scientific Management (see e.g., Brannick, Levine, & Morgeson, 2007; Fleishman & Reilly, 1992; Rasmussen, Pejtersen, & Goodstein, 1994; Shippmann et al., 2000; Vicente, 1999, for overviews). Similar to the application-oriented constructs, the point of departure are specific work environments and situations; more specifically the very tasks that are becoming less routine and more complex (i.e., task-oriented job analysis, e.g., hierarchical task analysis; Annett & Duncan, 1967; Shepherd, 2001) or the individuals performing these tasks and required to cope with the changes in work (i.e., worker-oriented job analysis, e.g., position analysis questionnaire; McCormick, Jeanneret, & Mecham, 1972; see Clifford, 1994; Dierdorff & Wilson, 2003; Levine, Ash, Hall, & Sistrunk, 1983; Pearlman, 1980, for comparisons).

But in contrast to the application-oriented constructs discussed above, the work and task analysis literature also has a long history within the domain of OP research and with it, a comprehensive nomological network, standardized and viable empirical paths to empirical insights, as well as tested relations to established (psychological) constructs. For example, the already mentioned U.S. Department of Labor's O*NET, offers a comprehensive overview on all kinds of occupations, jobs, and work tasks and the corresponding requirements by building on work and job analysis as well as established models of human functioning in work environments (N. G. Peterson & American Psychological Association, 1999). Building on this collection of insights, work and job analysis also enables the construction of reliable and valid assessment instruments tailored to the requirements of specific situations or occupations, overcoming one of the problems of application-oriented constructs (e.g., Fleishman & Reilly, 1992).

Generally speaking, coping with the changes in the world of work via job and work analysis allows for a bottom-up approach, for example quantifying the changes in requirements across work tasks (e.g., Cascio, 1995) or describing and comparing newly emerging jobs and their requirements within established frames of reference (e.g., Naikar et al., 2006; Vicente, 1999). This way, newly emerging jobs and requirements can be evaluated with the help of established approaches to empirical measurement, theoretically integrated and validated constructs, and compared to a large collection of already available data. In summary it therefore seems like a straightforward idea to cope with the changes in the work environment via job and work analysis. However, Neubert, Mainert et al. (2015) also identify several drawbacks of this approach: The need for detailed specification, the focus on already existing jobs and requirements, and the amount of necessary resources.

Firstly, while job and work analysis can rely on an established set of approaches, both methodologically and with regard to relevant content, the very nature of non-routine tasks leads to conceptual problems when trying to specify work tasks and requirements, which becomes even more prevalent when looking at problem solving in situations as described by Funke et al. (in preparation): It is difficult to describe a job whose content and requirements are changing constantly in a standardized way or to incorporate the requirements resulting from problem situations with open outcome criteria and unfamiliar methods for solutions into standardized questionnaires and performance measures (but see Braune & Foshay, 1983; Naikar et al., 2006; Shippmann et al., 2000).

Similarly, Neubert, Mainert, et al. (2015) point to the problem of focusing on already existing jobs and occupations, when looking for ways to deal with the requirements of future work environments whose jobs and occupations are not yet available for analysis (but see Schneider & Konz, 1989). This aspect becomes even more important when looking at the point of departure of the application-oriented constructs presented above: The need to cope with the developments of the working world proactively seems to be a strategic necessity for organizations (e.g., Shippmann et al., 2000) and an approach rooted in the analysis of already existing jobs and occupations naturally has limits in this respect (but see Cascio, 1998; Harvey & Bowin, 1996; Siddique, 2004, for attempts at integrating job and work analysis with a strategic view on human resource management).

Finally, coping with the changes in the work environment via job and work analysis also requires considerable investments in terms of time and resources (e.g., Levine, Sistrunk, McNutt, & Gael, 1988). New jobs, occupations, work tasks, and requirements need to be identified, measured and analyzed, which requires effort and skilled analysts, an investment that might be even less warranted once they change regularly or the strategic contribution of the analysis remains unclear (see also above). The problem can be somewhat mitigated by relying on existing resources, such as the O*NET (e.g., McEntire, Dailey, Osburn, & Mumford, 2006), but even then the necessary investments remain rather high, especially for small and medium sized organizations.

Taken together, Neubert, Mainert et al. (2015) point towards several problems of reacting to the changes in the world of work via job and work analysis alone, although the approach circumvents problems related to the nomological network and measurement instruments as major obstacles for application-oriented constructs. Interestingly, some of these problems might be handled by incorporating insights from CPS research into the established frame of reference given in job and work analysis. For example, the problem of specific investments for each and every new job might be mitigated by the identification of requirements of complex problem situations spanning across jobs and work tasks that have been already linked to more general trends in the work environment (e.g., the rise of problems requiring the handling of conflicting goals). The availability of CPS in the O*NET offers a way forward in that respect, although conceptual and methodological questions will have to be answered before integrating both approaches more closely. Similarly, the general trends in the work environment combined with the insights generated in CPS research might also allow for the estimation of future requirements in domains developing towards more complex problems.

Coming back to CPS research, the approach of job and work analysis actually looks quite interesting: The availability of a comprehensive approach describing jobs, work tasks, and requirements connected to established constructs facilitates the creation of more comprehensive models of CPS competency, thereby broadening the picture on CPS in general. For example, the KSAO model of CPS competency introduced by Fischer and Neubert (2015) directly builds on the same distinction from the work and task analysis literature and the models developed there.

Similarly, the approaches to describe and quantify work tasks and requirements in work environments developed in job and work analysis might offer a way to get CPS research closer to the actual problem solving processes in work environments, both methodologically (e.g., via work analysts or the assessment of a broader set of skills) and content-wise (e.g., taking into account different domains of expertise or additional constructs, such as systemic thinking, e.g., Rouwette & Vennix, 2006). Finally, encompassing collections of work tasks and requirements, such as the mentioned O*NET allow for the incorporation of CPS into an established frame of reference when evaluating its importance for specific jobs and work tasks: The relative importance of dealing successfully with complex problems can be compared to the handling of social interactions or the need for manual dexterity. A closer look at the potential for CPS research of incorporating insights from job and work analysis will be also taken in Chapter 5.

The third approach discussed by Neubert, Mainert, et al. (2015) is dealing with the changes in the work environment and the trend towards more complex and non-routine problems via the use of established psychological constructs, such as intelligence or personality. While the relation of CPS to these constructs has been highlighted in Chapter 2 (see also e.g., Greiff & Neubert, 2014; Kretzschmar et al., 2016; Süß, 1996), the constructs also have a long and successful tradition in OP (e.g., Barrick & Mount, 1991; Hunter & Schmidt, 1996; Schmidt & Hunter, 1998; Schmidt et al., 1988; Schooler et al., 1999). Established psychological constructs, such as intelligence can be seen as precursors of performance across very different situations, including work (e.g., Gottfredson, 1997b), have been shown to be of overarching importance across different occupations (e.g., Hunt & Madhyastha, 2012; Judge et al., 2004; Schmidt & Hunter, 2004), and hence, have been also used to investigate answers to the changing world of work towards non-routine and complex problems (e.g., Scherbaum, Goldstein, Yusko, Ryan, & Hanges, 2012). Contrasting to application-oriented approaches, the long tradition of researching basic constructs within psychology leads to comprehensive integration and differentiation with other constructs and reliable and valid assessment instruments. And in contrast to job and work analysis, the basic constructs are not bound to specific occupations or work tasks or the detailed analysis of specific jobs.

Nonetheless, Neubert, Mainert, et al. (2015) also identify difficulties when coping with the changes in the work environment via basic psychological constructs alone. Looking at specific complex and non-routine problem situations and the resulting requirements for individuals, the combination of all potentially relevant constructs becomes a complex problem itself. Consequently, straightforward answers to the changes of the work environment are hard to find (e.g., establishing strategies for organizations), a problem that is also underlying calls for a clarification of the paths between basic psychological constructs, such as intelligence and personality and actual OP-related behavior (Brouwers & Van De Vijver, 2012; Oswald & Hough, 2012). When looking at research focusing on the basic psychological constructs themselves, the disparity between actual decision processes and basic psychological research has also fueled the calls to take real-life decision processes more seriously in (cognitive) research (e.g., in NDM, see G. Klein, Orasanu, Calderwood, & Zsambok, 1993).

Additionally, there are also requirements resulting from the changes in the work environment, such as the need to actively explore and monitor the interplay of multiple influencing factors, the need to cope with one's emotions during problem solving, or the need to elaborate and balance multiple conflicting goals that are not included in basic psychological constructs in a straightforward way (e.g., J. Funke, 2010). This lack of including features and requirements of complex and non-routine problems into the assessment of basic psychological constructs becomes even more relevant when taking their importance for the work environment into account (as the research from NDM has impressively shown, e.g., G. Klein, 2008; Zsambok & Klein, 1997).

Compared to the previous approaches of application-oriented constructs and the use of job and work analysis, the interrelation of CPS with basic psychological constructs, such as intelligence and personality has been the focus of considerable research efforts (see Chapter 2). Interestingly enough, similar problems as highlighted for the basic psychological constructs above, such as the call for an integration of more heterogeneous aspects and actual problem solving processes have been also raised for CPS research itself, (e.g., Fischer & Neubert, 2015; J. Funke et al., in preparation). So while the consideration of complex and non-routine problem environments lies at the heart of CPS research (e.g., J. Funke, 2010), the connection to actual problem solving efforts in the domain of OP might require similar efforts as identified for basic psychological constructs.

4.2.2. Exploring Areas of Application

To facilitate this exploration of the role of CPS in the domain of OP, Neubert, Mainert et al. (2015) identify potential consequences and insights for researchers in several areas of application within OP: Personnel selection, career development, and organizational change and leadership. In personnel selection, Neubert, Mainert et al. (2015) see a potential role for CPS in the context of achieving congruency between the attributes of an individual and the work environment, the so-called person-organization fit (P-O fit, Chan, 1996; see Caplan, 1987; Edwards, 1991; Kristof, 1996, for reviews). P-O fit has been linked to organizationally relevant outcomes such as commitment and turnover, highlighting it's practical relevance (e.g., Adkins, Russell, & Werbel, 1994; Chatman, 1989; O'Reilly, Chatman, & Caldwell, 1991; Rynes & Gerhart, 1990). A potential role for CPS can be seen in offering an integrative perspective towards the different elements leading to a high fit between an individual required to handle complex and non-routine problems and an organization in need of capable members (i.e., integrating meta-cognitive skills and emotion regulation, see J. Funke, 2010). From the perspective of CPS research, interesting questions arise when looking at the role of organizational environments and practices influencing the process of individual problem solving in non-routine and complex problem situations (Foss & Weber, 2016; Nickerson & Zenger, 2004; see also the organizational literature on task complexity, Hærem et al., 2015; Pentland & Hærem, 2015, which might serve as a point of departure for further integration). In this view, the adaptability of problem solving efforts to different organizational environments and thus, the context dependence of adaptive problem solving strategies should lead to interesting insights for CPS research.

Similarly, the trend in the career literature towards new career paradigms, emphasizing the transferability of skills, the role of lateral mobility, and the need of adapting to different roles and organizations offers promising avenues for collaboration (e.g., M. B. Arthur & Rousseau, 1996; Bird, 1994; Greenhaus, Callanan, & Kaplan, 1995; Hall & Mirvis, 1996). The rising importance of so-called transversal skills and transferable knowledge in the career literature (Anakwe, Hall, & Schor, 2000; P. Griffin et al., 2012) has overlaps to similar discussions in education, where the findings from CPS are already receiving some attention (e.g., P. Griffin et al., 2012; Neubert et al., 2017; OECD, 2014). CPS might indeed serve as an indicator of different elements allowing individuals to cope with complex and non-routine problem situations across different work environments or integrating with different professional knowledge structures (see Fischer & Neubert, 2015; Neubert et al., 2017). And as in the case of personnel selection, important prerequisites for successfully handling complex problems throughout an individual's career can be highlighted and strengthened via CPS.

Again, the collaboration with researchers and practitioners from OP and career research might allow for interesting insights within the domain of CPS research as well. The different trajectories of an individuals' capabilities in handling complex problems are only beginning to surface (e.g., Frischkorn, Greiff, & Wüstenberg, 2014) and the question of transferability between different settings or in this case jobs, occupations, or work tasks is certainly an interesting avenue for CPS research, given the history of investigations into transferability within the domain of cognitive research (e.g., Gick & Holyoak, 1983; Holyoak, 2005; Reed, 2012). The question of transferability becomes even more prevalent, when taking into account recent models of CPS competency incorporating different aspects of complex problems and their requirements (e.g., Fischer & Neubert, 2015; J. Funke et al., in preparation), as well as the high relevance of (knowledge) transfer for organizations (e.g., Ashworth, 2006; Osterloh & Frey, 2000; Szulanski, 2000; Szulanski, Cappetta, & Jensen, 2004). It remains to be seen, whether the different prerequisites to successfully handling complex problems highlighted for example in the KSAO-model of CPS competency (Fischer & Neubert, 2015) can be transferred equally well to new contexts and problems and the investigation of career patterns might offer valuable insight in this regard.

The examples of P-O fit and the career literature also point towards potential roles of CPS for HR development and learning on a strategic level. The strategic importance of having human resources capable of dealing with complex problems has been highlighted recently (e.g., Metcalf & Benn, 2012; Raffiee & Coff, 2016), and viewing CPS competency as an (additional) indicator of Human Capital (HC) therefore promises an indicator of this strategic resource detailing the circumstances and boundaries of transfer between different settings (Gathmann & Schönberg, 2010). Neubert, Mainert, et al. (2015) point towards the notion of detailing HC beyond firm specific knowledge and general abilities as usually handled in HC theory (see also e.g., Yamaguchi, 2012). For CPS research, the link to the strategy and HC literature offers new views on the importance of the research topic and highlights the need to develop practically applicable concepts, instruments, and research findings. Also, the question of whether and how CPS competency can be trained becomes even more important given these contexts and should focus future efforts in this regard.

The fourth and final exploration of combining CPS and OP in Neubert, Mainert, et al. (2015) considers the case of leadership and the potential role for CPS in that area of application. Neubert, Mainert, et al. (2015) highlight previous research efforts highlighting the role of leadership in enabling employees to deal successfully with problems (Reiter-Palmon & Illies, 2004) as well as efforts dedicated to the problem solving of the leader (Mumford, Zaccaro, Harding, Jacobs, & Fleishman, 2000; Zaccaro et al., 1997). Interestingly, there seems to be quite some overlap with regard to important elements, such as problem construction, information encoding, and implementation and monitoring in CPS (e.g., Fischer et al., 2012) and leadership research (e.g., Mumford, Mobley, Reiter-Palmon, Uhlman, & Doares, 1991).

With regard to OP and leadership research, CPS research might provide hints at differentiating complex problems and the corresponding competency models. At least as interesting are the potential insights generated for CPS research. Leadership research has shown the importance of accounting for the social and collaborative side of problem solving, emphasizing social context and social skills for successfully dealing with complex and non-routine problems (Guarana & Hernandez, 2015; Shotter & Tsoukas, 2014; Zaccaro, Mumford, Connelly, Marks, & Gilbert, 2000). And while recent developments in the area of Collaborative Problem Solving are taking these aspects seriously as well (Greiff, Holt, & Funke, 2013), the impact of social elements on individuals handling complex problems certainly deserves more attention (e.g., Hesse, Care, Buder, Sassenberg, & Griffin, 2015; Hutchins, 1996). The insights generated in leadership research might provide blueprints for CPS research in this regard, integrating the social side of problem solving in applied contexts.

In summary, the paper by Neubert, Mainert, et al. (2015) provided a discussion of the intersection between research dedicated to CPS and OP. Alternatives for coping with the changing work environment from the domain of OP were discussed and areas of application were explored. Both, alternatives and areas of application already show the potential for further mutual learning, for example by strengthening the link between actual work tasks and CPS. As the article was meant to start a discussion with researchers and practitioners from OP, the corresponding comments published in the same issue of *Industrial and Organizational Psychology* are an at least as important part of the exchange that will be discussed in the following.

4.3. Starting the discussion with researchers and practitioners from OP: The comments on Neubert, Mainert et al. (2015)

The article by Neubert, Mainert, et al. (2015) was written from the perspective of CPS research with the organizational domain of application in mind. One of the noteworthy features of the journal the article was published in, *Industrial and Organizational Psychology*, is the possibility for researchers and practitioners from OP to publish comments to focal articles. As such, the comments by De Fruyt, Wille, and John (2015), Morelli, Illingworth, and Handler (2015), Oh (2015), Riggio and Saggi (2015), Sliter (2015), Su, Golubovich, and Robbins (2015), and Varghese, Lindeman, and Santuzzi (2015) present a first reaction from the domain of OP to the thoughts and proposals laid out in Neubert, Mainert, et al. (2015) and should provide some valuable input on future directions of inquiry and inter-disciplinary learning.

Generally speaking, there are three major directions visible in the comments with high relevance for CPS research: (1) Agreement with the observation of a general trend towards non-routine and complex problems in the world of work, (2) proposals for an expansion of the broader picture towards competencies and other important 21st century skills, and (3) critical remarks towards the general utility of introducing CPS to the domain of OP. With regard to the general trends in the world of work towards non-routine and complex problems, almost all authors agree with the analysis in Neubert, Mainert, et al. (2015). For example, the comment by Morelli et al. (2015) emphasizes the need to deal with the changes proactively from a practitioners point of view in a “future facing orientation” (Morelli et al., 2015, p. 269f). Similarly, De Fruyt et al. (2015) point towards the need of “connecting differential psychologists’ models of human differences and functioning with human resources professionals’ interest in understanding and predicting behavior at work” (De Fruyt et al., 2015, p. 276f) in light of the changes in work from a researcher’s perspective. Generally speaking, these assertions underline the importance of the general topic for both sides of the exchange between CPS research and OP on a general level and fall in line with a consideration of the strategic aspects of handling complex problems in the CPS literature. But while the general trend towards non-routine and complex problems is acknowledged by De Fruyt et al. (2015), they also point towards the second general trend throughout the comments, the need to take the broader picture into account, and specifically the literature on competency modeling.

Similar to De Fruyt et al. (2015), the comment by Sliter (2015) agrees with Neubert, Mainert et al. (2015) with regard to the changes in the world of work, while highlighting the utility of competence-based models to cope with these changes based on currently available models, tools, and findings from OP. And while De Fruyt et al. (2015) argue for the inclusion of more constructs and broader taxonomies from the network of 21st century skills in OP under the umbrella of competency modeling, Sliter (2015) points to the practical utility and the long and successful history of already established competency models in the tradition of McClelland (1973) in OP. Specifically, she highlights recent advancements in competency modeling by Campion et al. (2011) merging different models into a comprehensive KSAO model and the successful handling of non-routine and complex problems and their requirements on the basis of competency models in OP (Rodriguez, Patel, Bright, Gregory, & Gowing, 2002).

Similar trends towards incorporating the notion of competencies and the accompanying conceptions and research findings have surfaced within CPS research recently (Fischer & Neubert, 2015; J. Funke et al., in preparation), proving the potential for insights through interdisciplinary exchange. Specifically, the model introduced by Fischer and Neubert (2015) is utilizing the very differentiation of knowledge, skills, abilities and other components brought together by Campion et al.

(2011) in the domain of OP to get to a more encompassing picture of human interaction with complex problems in the domain of CPS research. It remains an open question, whether competency models developed in the area of CPS develop similar potential for additional insights in the domain of OP.

One important point that is raised by Sliter (2015) is the possibility to handle the needs resulting from the changes in the world of work via competency models building on already established psychological constructs, such as intelligence, personality, and interests. This skepticism towards the unique contribution of CPS compared to already established constructs is also mirrored in the comments by Su et al. (2015), Varghese et al. (2015), and Oh (2015). Specifically, Su et al. (2015) point towards the underlying conceptual problems in some models of 21st century skills and ask for a critical evaluation of new constructs in this regard. Following an evaluation by the National Research Council (NRC), they view 21st century skills (including CPS) as a combination of already established elements, rather than new constructs altogether (National Research Council, 2012). From the perspective of CPS research, this comment highlights the value of efforts, such as those by Kretzschmar, Neubert et al. (2016) or Greiff and Neubert (2014), differentiating CPS from basic constructs, such as intelligence or personality (see Chapter 2).

An interesting direction of thought following a similar line of reasoning is put forward by Varghese et al. (2015), who highlight the potential and need to account for mediating factors when exploring the relation between constructs, such as intelligence and personality and performance indicators in OP. Specifically, they point towards the findings on the role of job skills for the relation of intelligence and job performance and the role of situational factors in determining the effect of personality traits in organizational environments (Hunter, 1986; Motowildo, Borman, & Schmit, 1997). In terms of CPS research, the question of mediating factors corresponds with the notion of 'complex cognition' as a combination of different underlying factors necessary for the handling of complex problems as raised by Funke (2010) and competency models of CPS incorporating different underlying factors including those raised by Varghese et al. (2015). The debates on the construct validity of different assessment approaches within CPS research also point in the direction of further need for clarification on the role and positioning of CPS in the larger picture of (non-)cognitive skills and abilities (see J. Funke, 2014a; Greiff & Martin, 2014; Schoppek & Fischer, 2015).

Similarly related to a clarification of alternatives, Oh (2015) points towards the need to compare the contribution of CPS research to established tools from the area of Assessment Center (AC) research that already includes very similar content on a descriptive level while also drawing together contributions from basic psychological constructs. For example, she mentions the dimension “problem solving” given within the AC model presented by Arthur, Day, McNelly, and Edens (2003) that has a definition with considerable overlap to the elements included in models on CPS (cf. the dimensions described in Dilchert & Ones, 2009; Fischer et al., 2012). And indeed, there seems to be lot of opportunities for researchers from the domain of CPS to learn from the area of AC research on the combination of different constructs within a single session of assessment, especially given the competency-oriented models of CPS.

Importantly, previous efforts in the domain of CPS research, such as those presented in Chapter 2, have already established the close connection, but also distinctness of CPS from other factors of intelligence. Even more, first explorations into the utility of microworlds developed in CPS research in personnel selection have highlighted the beneficial features in terms of acceptance by applicants as a central advantage of microworlds compared to standard tests of intelligence (e.g., Kersting, 1999). But as mentioned above, the need to clarify the relation and positioning of CPS within these nomological networks remains, especially in light of argument-based models of validity (e.g., Kane, 2001, see Chapter 2.2).

In addition to highlighting the role and importance of basic psychological constructs for performance in the domain of OP, Varghese et al. (2015) also notes models accounting for the situational influences on job performance, such as those proposed by Tett and Burnett (2003). For example, the influence of personality factors has been shown to depend on the specific situation an employee is in; Varghese et al. (2015) mention the varying influence of a high need of autonomy depending on the supervision ranging from thriving under democratic leadership to withering under authoritarian leadership. With regard to CPS research, this observation of the context dependence of the effects of specific behaviors or traits can be also observed for the case of strategies. Fischer and Neubert (2015) point to the example of the VOTAT strategy being applicable in a broad range of scientific settings (e.g., Scherer & Tiemann, 2012), while exhibiting detrimental effects in situations of interpersonal conflicts. Models such as those put forward by Tett and

Burnett (2003) might serve as a guideline of how to include situational factors into the examination of CPS (for a broader discussion of the need to include contextual factors see e.g., Johns, 2001).

In summary, the comments published in reaction to the focal article by Neubert, Mainert, et al. (2015) offer some valuable insights on potentially relevant developments for CPS research, thereby completing the discussion of the potential benefits of CPS assessment from the perspective of OP. Among the most important aspects raised by the comments are certainly the confirmation of the general trend towards non-routine and complex problems also in the domain of OP and the acknowledgement by some commenters, that CPS might indeed offer a valuable perspective in this regard. Furthermore, a number of comments reference competency models as an important framework within the domain of OP that might accommodate the investigation of the broad range of constructs relevant for the handling of complex problems in actual work situations. Some of these observations have already found their way into CPS research, such as the introduction of competency models, while others require further attention, such as the calls for an incorporation of a broader set of influencing factors mediating the influence from basic psychological constructs to job performance or the consideration of situational influences determining the effect of specific behaviors or traits. But at least as important as the development of future avenues for research is the beginning of an interdisciplinary discussion between the domains of CPS research and OP.

5. Discussion and outlook

Relating the domain of CPS research to research in OP was the overarching goal of this dissertation. Following the general observation of a world developing towards complex problems and a short look at the history of CPS, inquiries were structured around three strands directed at the construct of CPS and its nomological network (see Chapter 2 and Kretzschmar, Neubert, et al., 2016), the construct's assessment with the help of finite state automata (Chapter 3 and Neubert, Kretzschmar, et al., 2015), and finally the connection to research and practice in OP (Chapter 4 and Neubert, Mainert, et al., 2015).

Supporting the core papers of the dissertation, a range of additional efforts was undertaken, integrating a broader range of constructs into the nomological network of CPS, specifically ICT literacy and personality (Greiff et al., 2014; Greiff & Neubert, 2014). Furthermore, in addition to the discussion of the potential of CPS research for OP, the utility of CPS was also explored within other domains of application, namely education in schools, vocational education, lifelong learning and education, and career research (Greiff, Neubert, Niepel, & Ederer, 2015; Kretzschmar et al., 2014; Mainert, Kretzschmar, Neubert, & Greiff, 2015; Neubert et al., 2017). Finally, an integrative theoretical position was developed, the KSAO model of CPS competency, presenting a first attempt at a competency-based model building on established differentiations in OP (Fischer & Neubert, 2015). In the following, the contributions provided in the different strands of inquiry will be discussed from a broader perspective, including the evaluation of critical remarks and limitations and an outlook on the potential for future research efforts will be given.

5.1. The construct of CPS: Towards a comprehensive nomological network

With regard to the construct of CPS, the clarification of the relation to intelligence in Kretzschmar, Neubert et al. (2016), as well as those directed at additional (psychological) constructs in Greiff and Neubert (2014) and Greiff, Kretzschmar, et al. (2014) seem like a valuable contribution to the nomological network of CPS. Specifically for the case of intelligence, the work presented in Kretzschmar, Neubert et al. (2016) and the empirical evaluation of the redundancy and distinctness perspective represents an important extension of previous research, especially in light of the overarching importance of intelligence in OP (see also the critical remarks by Oh, 2015, which fall into line with the redundancy perspective).

Specifically, the article went beyond existing research in combining the MCS approach to CPS assessment with a broad operationalization of intelligence, while also accounting for external criteria and different measurement models. Thereby, strengths of previous efforts from the two perspectives that could explain differences in previous results were integrated.

The results of Kretzschmar, Neubert et al. (2016) give support to the distinctness perspective, in highlighting distinct and stable variance not captured in broad operationalizations of intelligence, although the close connection between intelligence and CPS is undeniable (see also Kretzschmar, 2015). Thereby, the findings presented in the first core paper of the dissertation relate to one of the fundamental problems in CPS research in clarifying the role and relation of CPS to intelligence. Similar to the findings in the article by Kretzschmar, Neubert et al. (2016), the two additional papers targeting personality and ICT literacy found evidence for distinct features of CPS, compared to the established constructs (see Greiff et al., 2014; Greiff & Neubert, 2014).

From the perspective of OP and the close relation between intelligence and CPS in Kretzschmar, Neubert et al. (2016), it seems important to note that sub-facets of intelligence with a similarly strong relation to the overarching construct have historically been lacking in additional predictive power with regard to external criteria in OP (e.g., Brown, Le, & Schmidt, 2006; Lang et al., 2010). This finding might question the claims of a higher utility of a dedicated CPS assessment compared to intelligence when focusing on predictive power alone. Interestingly, the results for the external criterion included in Kretzschmar, Neubert et al. (2016) point towards similar problems, that is, a lack of explaining additional variance through CPS compared to a broad operationalization of intelligence, although with regard to an external criterion from the educational domain. Furthermore, the close relation between CPS and intelligence identified in Kretzschmar, Neubert et al. (2016) also point towards the need to position CPS more clearly in its relation to intelligence, for example as a distinct sub-facet of intelligence in CHC-models of intelligence (e.g., McGrew, 2009) or as a distinct construct going beyond intelligence altogether (see also Kretzschmar, 2015).

Importantly, recent developments in CPS research towards competency models of CPS (Fischer & Neubert, 2015; J. Funke et al., in preparation) were not part of the considerations in Kretzschmar, Neubert et al. (2016) and it remains to be seen whether developments in this regard will lead to additional insights on the relation of CPS and intelligence in the future (see also Schoppek & Fischer, 2015). Positioning CPS as a sub-facet of intelligence versus a combination of different skills, abilities, knowledge and other components has important conceptual implications (J. Funke, 2010; see also Kluge, 2008b; Tricot & Sweller, 2014).

For example, the question of malleability and trainability should be strongly influenced by the view taken on the importance and positioning of CPS and intelligence. That is, if performance in CPS is seen as being largely determined by intelligence, interventions targeted at the general capability of individuals to handle complex problems successfully are facing obstacles similar to those directed at strengthening reasoning ability or other sub-facets of intelligence (e.g., Klauer & Phye, 2008; see also the discussion of application-oriented constructs in Neubert, Mainert, et al., 2015). In contrast, if CPS competency is seen as a combination of knowledge, skills, abilities and other components sensu Fischer and Neubert (2015), training interventions can be targeted at elements that problem solvers are lacking in different complex problem situations and that are known to be malleable, for example explicit knowledge structures.

Similarly, the importance of psychometric criteria and the associated evaluations of constructs and their measurement has been questioned by researchers focusing on competency models in light of training and development (e.g., Sanchez & Levine, 2009), highlighting differences in goals along different dimensions.[28] From this perspective, the strong focus on psychometric features in recent CPS assessment might actually be counter-productive for an application in OP via competency models of CPS, as it obstructs straightforward pathways to training and development.

In summary, interesting implications are resulting from a positioning of CPS and intelligence in one way or the other, thereby increasing the importance of reflecting on the choice of approach and modeling taken. The results in the first core article of this dissertation are only a first step towards systematically embedding CPS in this respect (see also Kretzschmar, 2015). CPS research might actually benefit

[28] For example, the purpose of assessment being description vs. influence, the focus on the enactment of roles vs. objective description, an orientation towards past or future performance and so on (e.g., Sanchez & Levine, 2009).

from the collaboration with the domain of OP through an explication of perspectives and associated assumptions in an applied setting.

Comparing the specific focus on individual problem solvers in current CPS assessment to theory and empirical research in OP might also point towards areas of future development for CPS research at large. As the comments on Neubert, Mainert et al. (2015) show, there are also different positions within OP concerning the importance and role of ‘classical’ psychological constructs. For example, Varghese et al. (2015) are highlighting the importance of established constructs for assessment and theory, while also pointing towards the need to understand intermediating factors and their impact in specific circumstances. What seems interesting in that regard is the potential arising from a more direct application of CPS in specific domains, connecting research on CPS closer to actual needs of practitioners in the respective field. In this sense, the first core paper of the dissertation represents a contribution to one perspective of exploring the interrelation between CPS and intelligence, while also pointing to the need to consider the relation in more detail and from additional theoretical and empirical perspectives.

Specifically related to CPS, given recent disputes on adequate approaches to the measurement of CPS (J. Funke, 2014a; Greiff & Martin, 2014), an explication of differences in underlying assumptions might offer remedy in explaining different positions while also offering a way forward (see also Schoppek & Fischer, 2015). At the core of the debate, there seem to be differences, into what could and should be regarded as complex problem, how to conceptualize the construct of CPS, and how to identify CPS competency in an assessment situation (see also below). Similarly, Hafenbrädl, Waeger, Marewski, and Gigerenzer (2016) have highlighted the importance of reflecting on the conceptual lens when looking at human behavior, especially in applied settings, by translating ideas brought forward by Allison (1969) to the area of human problem solving. In their view, the explication of a framework of assumptions via a conceptual lens leads from description to explanation, prediction, and prescription.[29] This development towards contributing to a better handling of complex problems by individual problem solvers seems to be a worthwhile undertaking in the case CPS in a world developing towards complexity.

[29] Interestingly, this view also aligns with an argument-based view on validity (e.g., Chapelle et al., 2010; Kane, 2001, 2006), see also Chapter 5.2.

In this regard, the contributions to understanding the relation of CPS to established (psychological) constructs, such as intelligence is an important extension of a specific perspective on CPS, especially given the high relevance of the construct in the domain of OP, while also requiring further efforts in the future, especially in terms of explicating the underlying conceptual lens and the associated framework of assumptions (see also Chapter 5.3).[30] On the other hand, there is also need for additional explorations into the actual translation of constructs into empirical investigations, and hence, the need for research into the assessment side of CPS.

5.2. The assessment of CPS: Combining finite state automata and multiple complex systems.

The second core paper of the dissertation is even closer related to current application of CPS and more specifically, the development and evaluation of specific measurement instruments (see also Fischer, 2015; Kretzschmar, 2015). As laid out in Chapter 3, the combination of MCS and FSA facilitates the development of CPS instruments going beyond the strong focus on VOTAT as found in previous instruments building on MCS. With the introduction and empirical evaluation of a set of MICROFIN tasks, Neubert, Kretzschmar et al. (2015) have shown the possibility of extending the measurement of CPS to exploration beyond VOTAT within the framework of MCS, while retaining the advantages of the MCS-approach, such as multiple independent indicators. Even more, the empirical results presented in Neubert, Kretzschmar et al. (2015) also point towards the psychometric viability of this approach in terms of reliability and convergent validity in relation to already established instruments building on LSE. So in summary, the extension of current CPS assessment via the combination of finite state automata and multiple complex systems seems successful. Similarly, in view of an application in the domain of OP, the availability of a broad set of tasks targeting the construct seems worthwhile as well, as a valid and reliable assessment of the construct can be seen as a prerequisite to contributions in the domain (see e.g., Neubert, Mainert, et al., 2015).

Nevertheless, there are also aspects related to the assessment of CPS that are not as positively looking in comparison, especially when extending the view beyond the focus on current assessment-oriented approaches and the refinement of

---

[30] The work by Hærem, Pentland, and Miller (2015) in the area of task complexity might provide an interesting starting point in that respect, see Chapter 5.3.

instruments within the given formal frameworks. Specifically, there are problems related to the sampling of complex problems, the role of prior knowledge, and the general approach taken in current CPS assessment.

Importantly for an application in OP, the contribution by Neubert, Kretzschmar et al. (2015) is not building on a (representative) sampling of complex tasks and their requirements. Twenty years ago, Heinz-Martin Süß (1996) has highlighted the need to build CPS research and assessment on a more comprehensive overview on the tasks and requirements found in actual instances of complex problem solving and the issue has not been solved sufficiently to this date (see also Kluwe, Misiak, & Haider, 1989).

Specifically, Süß (1996) raised the possibility and need to work with an actual sampling of complex problems in real-life settings, contrasting to the ad-hoc creation of artificial problem simulations typically found in CPS research. In this respect, the development of MICROFIN represents an example of selecting and refining different elements of CPS assessment (see Kretzschmar, 2015, for more details), while the more general problem of selecting appropriate complex problems in the first place was not tackled systematically.

The interaction with researchers and practitioners from the domain of OP might provide valuable insights in this respect, as research from OP and more specifically, work and task analysis has a long and successful history of systematically sampling work tasks and job requirements in all sorts of occupations (see also Chapter 4.2). Overarching frameworks of work tasks, such as those available in the already mentioned Occupational Information Network (O*NET; N. G. Peterson & American Psychological Association, 1999, see www.onetcenter.org) can provide CPS research with a point of departure for future inquiries.

On the one hand, these frameworks provide an established frame of reference and methodological guidance for the evaluation and comparison of different problem situations embedded in a specific context (see also Süß, 1996). Recent developments towards comprehensive models of CPS competency (e.g., Fischer & Neubert, 2015) have already begun to incorporate the theoretical frameworks differentiating and describing work situations for the benefit of CPS research and a utilization of the methods used in work and task analysis seems straightforward.

On the other hand, the data already being collected within frameworks, such as the O*NET might also provide insights for CPS research more directly. For example, the inclusion of 'complex problem solving' in the O*NET offers the possibility to look for occupations and tasks where work analysts see heightened importance to handle novel, ill-defined problems and use these occupations as a starting point for the sampling of complex problems in the world of work. Methodological differences with regard to the measurement of requirements and the identifications of interindividual differences make a direct transfer difficult, as current approaches to CPS assessment feature a performance-based rating of CPS skills compared to the rating of work tasks by a third person included in the O*NET. Nevertheless, the evaluation of requirements from the perspective of work analysis and the respective data collected there might provide useful empirical hints for an initial sampling of the type of tasks and problems individuals face when being confronted with complex problems in the domain of OP.

Future research will have to tell, whether a closer connection to research and practice in OP can help in overcoming the problem of sampling an adequate set of complex problems for CPS assessment. Interestingly, the emphasis on different contexts of application arising from the interaction with research in OP corresponds to the demand from argument-based validity approaches, calling for greater emphasis on the domain of application when evaluating the validity of measurement instruments (e.g., Chapelle et al., 2010; Kane, 2006). Historical instances of research closely connected to both, research in CPS and OP, such as the efforts by Wallach (1998) or Ackerman (1992) give hope for fruitful interaction between the domains in this respect.

A problem of similar importance to the quest of sampling representative complex problems that is also related to the connection of CPS assessment to the 'outside world' and that has not been tackled with the introduction of MICROFIN is the role and importance of prior knowledge and experience for the handling of complex problems within the framework of MCS. A range of researchers from different domains of psychology have established the overarching importance of prior knowledge for subsequent knowledge acquisition and other problem solving processes (see e.g., Dunbar, 1998; J. Funke, 1992; Kuhn et al., 1995, for first overviews).

Similarly, earlier instances of CPS research working with larger computer simulations have invested considerable time and effort into the clarification and assessment of knowledge during complex problem solving (e.g., Elio & Scharf, 1990;

J. Funke, 1992; Kersting, 2001; Kersting & Süß, 1995; Süß, 1996). In contrast, CPS research building on the MCS approach considers typical instruments employed in CPS assessment to be 'knowledge-lean' (J. Funke, 2014a; see also Quesada et al., 2005), thereby largely neglecting these insights on the role and importance of knowledge for problem solving features and processes in specific problem solving environments (e.g., Roth, 1998; but see e.g., Beckmann & Goode, 2013; McElhaney & Linn, 2011) as well as the complex relation between understanding a problems components and relations and selecting fit explorations strategies (e.g., Gaschler, 2009; Levy & Wilensky, 2011).

Future investigations will have to tell, whether the MCS approach can be also extended to problems that require a deeper understanding of different fields, and therefore, take the knowledge requirements of actual complex problems more seriously in assessment. Taking a more positive view, the combination of different problem solving scenarios within the MCS approach might also facilitate the exploration of this interplay between specific knowledge structures and problem solving processes (see e.g., J. Funke, 2008; Wittmann & Hattrup, 2004). Recent explorations in the domain of vocational education give hope for a fruitful interaction with the domain of OP in the future (see Neubert et al., 2017).

On a more general level, there are also more fundamental issues with the approach of MCS as realized in current assessment instruments targeting CPS besides the sampling of tasks and the role of prior knowledge: When looking at typical problems serving as a justification for a high relevance of CPS research, such as climate change, complex leadership problems or recent technological developments, all of them require quite comprehensive efforts in tackling the associated complex and dynamic challenges (e.g., Amelung & Funke, 2013; Geels, 2002, 2010; Geels & Schot, 2007; Kallerud et al., 2013; Mumford et al., 2000; H. C. Peterson, 2009; Vester, 2007; Wissenschaftsrat, 2015). In short, these problems cannot be handled via 'simple' domain-independent strategies (see also Ohlsson, 2012, and the critical reflections on general problem solving strategies there). For example, tackling complex leadership problems via the VOTAT strategy is headed for failure, as the (cost-free) return to a given situation is not possible in this type of situation (see already Rittel & Webber, 1973, and their thoughts on the impossibility to tackle social problems with scientific methods). Instead, these highly complex problems require, for example, dealing with ambivalence and different contextual interpretations (Guarana & Hernandez, 2015; Mumford et al., 2000).

Furthermore, according to Funke (2014a), the MCS-based instruments are built on a restricted conceptualization of CPS altogether, where participants can and should explore all potential relations exhaustively and have to deal with relatively well-defined problems (i.e., clear states, goals, and problem elements, see J. Funke et al., in preparation, for a discussion of different views on complexity). In this sense, the work by Neubert, Kretzschmar et al. (2015) is an extension of the MCS-approach while also suffering from the same drawbacks as established instruments building on the same framework (see already Ebbesen & Konečni, 1980).

In contrast, historical instances of problem simulations more closely aligned with research in OP and the task and work analysis literature, such as the POWERPLANT mentioned in Chapter 2 (see Wallach, 1998; Wallach & Tack, 1998) were built on a much closer connection between actual problem solving processes and assessment instruments (see also the efforts by Jonassen, 2004).[31] A closer dialogue between the research traditions in OP and CPS might be beneficial for future CPS assessment in this respect and enable a broader picture on assessment. For example, research in OP provides an established set of approaches to the analysis of specific problem situations (see above), which in turn might serve as a point of departure for the creation of problem solving scenarios aligned with specific complex problem situations (see also Neubert et al., 2017; and the call for more complex descriptions in OP by Tsoukas & Dooley, 2011).

Interestingly, some instances of CPS research also took a lot of interest in the details of specific problem solving processes, for example via detailed case studies or the fine-grained analysis of problem solving in specific computer simulated scenarios (e.g., Rollett, 2008). Similar explorations into actual processes of problem solving “in action” instead of prior conceptions of rational problem solving were performed in the domain of OP as well, for example the already mentioned studies targeting the problem solving efforts of managers (Metcalf & Benn, 2012; Mintzberg, 1973; Mumford et al., 2000; Wagner, 1991; Yukl, 2010). These studies exploring the

[31] Interestingly, the process of dealing with DYNAMIS-based tasks has been analyzed in detail by Rollett (2008) on multiple levels, ranging from single inputs to chains of interactions. And while the analysis seems applicable to MCS-based tasks building on LSE, such as MICRODYN or GENETICS LAB and even MICROFIN to some extent, systematic analyses have not been undertaken to my knowledge.

decision making of managers and other organizational members have repeatedly highlighted the heterogeneity and non-compliance with rational modes of problem solving; a finding that might very well benefit the understanding of actual individuals faced with complex problems in CPS research compared to the notion of 'just' comparing their exploration to VOTAT as an idealized mode of handling complex problems (e.g., Isenberg, 1986; Keegan & Kabanoff, 2007; Mezias & Starbuck, 2003; Mintzberg, 1973; Smith, 1997; Wagner, 1991).

In summary, the work by Neubert, Kretzschmar et al. (2015) can be certainly considered a contribution to CPS assessment building on the framework of MCS, refining and extending currently available measurement instruments. Nevertheless, the discussion and interaction with research and practitioners from OP has also highlighted a range of open questions concerning CPS assessment, some of them requiring further thought on a more fundamental level and inviting further explorations between both domains.

5.3. Complex Problem Solving and Organizational Psychology: Connecting discourses

The interaction and discussion with practitioners and researchers from the domain of OP lies at the heart of the third core paper of the dissertation. In contrast to earlier transfer between the domains of CPS research and OP, the developments in CPS research towards fleshing out the nomological network of the construct (see Chapter 2) or the development of psychometrically oriented assessment instruments (see Chapter 3) were not taken up in the discourse in OP on a broader scale. The focal article itself and the following comments from the domain of OP therefore represent an interesting exchange of ideas and insights between the domains of CPS research and OP (see also Neubert et al., 2017).

Specifically, Neubert, Mainert, et al. (2015) contrasted the development of assessment instruments building on MCS and formal frameworks in CPS research with alternative approaches already established in OP. By comparing CPS assessment to approaches building on application-oriented constructs, work and task analysis, and basic psychological constructs, the specific features and promises of CPS assessment were discussed and the potential of utilizing CPS assessment in a range of sub-domains of OP was explored. In the following, the comments featured in the same issue of *Industrial and Organizational Psychology* provide an interesting view from the perspective of OP on the thoughts presented in the focal article by Neubert, Mainert, et al. (2015). Interestingly, the general importance of dealing with complex problems for the domain of OP is emphasized throughout the comments, highlighting the relevance of the general topic for the domain of OP. At the same time, critical factors are raised by the practicioners and researchers that highlight the need of further efforts in the future.

The pointer towards competency models as a viable path to tackle the challenge of preparing individuals for complex problems proactively has already been taken up in recent theoretical considerations within the domain of CPS research (see Chapter 4), highlighting the benefit of crossing the boundaries of individual disciplines (see also e.g., Alvesson & Sandberg, 2014, for a more general discussion of this endeavour). Similarly, some of the critical accounts, such as the exploration of the intersection between assessment center research and CPS or accounting for situational factors could offer valuable opportunities for collaboration in the future (see Chapter 4.3).

Going beyond the points raised in the comments to Neubert, Mainert, et al. (2015), there are also interesting observations to be made when comparing the approach taken to the construct of CPS in CPS research to research in OP from a broader perspective (see also the requirement of explicating assumptions raised in Chapter 5.1). Building on a long tradition of (cognitive) research into individuals handling problems, CPS research is built around several assumptions on the important aspects to look for when researching CPS (see Chapters 1 and 2). Current CPS research thereby shares the assumptions with what Hærem et al. (2015) term “old assumptions” for the case of task complexity: Separability of task from behavior and context is assumed, focus is put on an individual level of analysis, predefined “types” or in the case of CPS features of complexity are analyzed, and a linear relation between task components and complexity is anticipated (see e.g., Stadler, Niepel, & Greiff, 2016, for a recent example).

Interestingly, this differentiation between assumptions in classical views on task complexity and contrasting newer developments, mirrors those between classical problem solving research and the developments summarized under the label of Naturalistic Decision Making (NDM, e.g., Ebbesen & Konečni, 1980; G. Klein, 2008; G. Klein et al., 1993; Lipshitz, Klein, Orasanu, & Salas, 2001). Similar to the case of task complexity in Hærem et al. (2015), researchers in NDM have tried to extend decision making research by building on a different set of assumptions, thereby embedding decision making research closer in the respective context of application. For example, NDM research has brought attention to factors influencing problem solving in specific real-life settings, such as perceived uncertainty, thereby blurring the line between task, behavior, and context (e.g., G. Klein et al., 1993; Lipshitz & Strauss, 1997; see G. Klein, 2008, for an overview on NDM). The relation between NDM and CPS has already been discussed in CPS research by some authors recently (e.g., J. Funke, 2010, 2014b; Schoppek & Fischer, 2015) and the interrelation of the construct from CPS research with an application in the domain of OP might stimulate the development of similar analytical progress for the case of CPS. Following the naming scheme in decision making research, these efforts could be organized under a label of “Naturalistic Problem Solving” or NPS. But how can one approach the effort of describing "Naturalistic Problem Solving"?

In OP, there is a long tradition of looking at one and the same problem from different perspectives (e.g., Volmerg et al., 1995) and in comparison, the (cognitively oriented) perspective on problems in CPS research could be extended by a range of features. For example, different conceptual lenses on complex problems have been proposed in the literature (see also J. Funke et al., in preparation), highlighting for example different features of complex, ill-defined, large-world, or wicked problems, each with their own and specific characteristics (see also Denison, Hooijberg, & Quinn, 1995). In comparison, current (empirical) CPS research is largely focused on the narrow set of features proposed by Buchner in Frensch and Funke (1995b), namely complexity, interconnectedness, intransparency, dynamics, and politely and their operationalization in rather technical ways (e.g., counting variables and relations Stadler et al., 2016, see also Chapter 2.1).

Approaches taking a more narrative route to the exploration of complex problems, such as those proposed by Jonassen (2004), Shore, Bernstein, and Lazer (2015) or Tsoukas and Hatch (2001), looking at very different groups of problems, the ambiguity inherently included in some complex problems, or including different perspectives in the description of singular cases, have been largely ignored in contrast (see also Dörner, 1989a). Additional candidates for potentially enriching perspectives would be recent developments in cognitive science summarized under the labels of situated (e.g., Robbins & Aydede, 2008; Roth, 2001), embodied (e.g., Wilson & Golonka, 2013), and distributed cognition (e.g., Hutchins, 1996; Zhang &

Norman, 1994) that have been already integrated in OP to some extent, but much less so in current discussions on CPS (e.g., Baron, 2014, see also the proposal of NPS above).

In OP in comparison, the heterogeneity of theoretical lenses of looking at complex problems has been integrated more closely into theoretical and empirical explorations (e.g., Aken, Berends, & Bij, 2010). For example, different perspectives on complexity have been discussed in OP, ranging from the detailed exploration of specific cases (e.g., van der Schaaf, 1993) to more complex theoretical models in work analysis (e.g., Hærem et al., 2015; Liu & Li, 2012; Wood, 1986). Similarly, researchers in OP have discussed different methodological approaches (e.g., Elqayam & Evans, 2011; K. J. Klein, Dansereau, & Hall, 1994; Sandberg & Tsoukas, 2011) and reflected on the role of the observer and environment in the determination of complexity and individuals handling them (e.g., Kämmer, Gaissmaier, & Czienskowski, 2013; Norman, 2013; Tsoukas & Hatch, 2001), together leading to the establishment of interdisciplinary research programs (e.g., P. Anderson, 1999; Axley & McMahon, 2006; Mainzer, 2009). And while some of these questions have been tackled in CPS research as well, the anchoring of the domain in a specific area of application has resulted in much more palpable outcomes.

Specifically, the question of whether research findings can be translated into relevant insights in practice has been tackled in the domain of OP in various ways (e.g., Kieser, Nicolai, & Seidl, 2015), even including the use of simulation methods for the purpose of training (e.g., Fisch & Beck, 2004; Wood, Beckmann, & Birney, 2009) or the influence of the work environment on the capacity to handle complex problems (e.g., Kohn & Schooler, 1978). CPS research might benefit from the experience gathered in OP, when trying to set up research that benefits actual problem solvers confronted with complex problems in addition to tackling problems related to the psychometric scaling of assessment instruments (see Chapter 4.3). On this rather superficial level, it seems like CPS research could actually learn a lot from considerations concerning the handling of complex problems in the domain of OP.

In summary, the interaction with researchers and practicioners from OP via the focal article by Neubert, Mainert, et al. (2015) and the following comments has raised a lot of interesting points when looking at the intersection of CPS research and OP (see also Neubert et al., 2017). Interestingly, the nomological network and thus, the relation to and differentiation from other (psychological) constructs was also raised as an important issue in the comments, as well as the need for reliable and valid assessment instruments (e.g., Oh, 2015; Su et al., 2015). So indeed, the efforts presented in the first and second core paper of the dissertation (see Chapters 2 and 3) seem to be of some relevance to researchers and practicioners from the domain of OP as well (see also Neubert et al., 2017). Additionally, the dialogue (fortunately) also raised a broad set of further questions and remarks that might serve as a point of departure for future exchange and interdisciplinary research.

6. Synopsis

The point of departure for the dissertation was the comparably low amount of interaction between current CPS research and an important area of application, the domain of OP. In light of overarching trends towards complexity in the world of work, as well as the lively exchange between a different area of application, namely that of education and current CPS research, this lack of transfer and translation seemed curios. Therefore, the dissertation was oriented towards three areas of importance when looking at the interaction between CPS research and OP, namely the nomological network of CPS, the availability of reliable and valid assessment instruments, and the exploration of the connection between CPS and OP in a discussion with researchers and practicioners from the domain of OP.

The efforts dedicated towards each of these three directions have resulted in very different scientific endeavours: The efforts presented in Chapters 2 to 4 have clarified the nomological network of CPS with regard to the distinction between CPS and intelligence and other important constructs, they have developed and validated a measurement instrument targeting CPS by building on finite state automata, and they have established a dialogue with researchers and practicioners from the domain of OP in *Industrial and Organizational Psychology.* Taken together, these contributions to each strand of inquiry have (hopefully) led to useful insights on the human handling of complex problem situations in their own right. Nevertheless, taking the broader picture of OP into account has also highlighted limitations present in current CPS research in each of these areas (see Chapters 5.1 to 5.3). So in summary, the connection between CPS research and the domain of OP has resulted in a range of interesting findings and maybe even more important, a set of additional perspectives and future research opportunities.

Each of these possiblities for further research, as well as the assemblage of pointers to additional exploration throughout this work promises interesting insights and innovation for the domains of CPS and OP in the future. Different doors of perception await their opening or in the words of Neubert, Mainert et al. (2015): Let's get started!

References

Abele, S., Greiff, S., Gschwendtner, T., Wüstenberg, S., Nickolaus, R., Nitzschke, A., & Funke, J. (2012). Dynamische Problemlösekompetenz: Ein bedeutsamer Prädiktor von Problemlöseleistungen in technischen Anforderungskontexten? [Dynamic problem solving - an important predictor of problem-solving performance in a technical domain?]. *Zeitschrift für Erziehungswissenschaft*, *15*(2), 363–391. https://doi.org/10.1007/s11618-012-0277-9

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: dynamics of ability determinants. *Journal of Applied Psychology*, *77*(5), 598–614. https://doi.org/10.1037/0021-9010.77.5.598

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227–257. https://doi.org/10.1016/S0160-2896(96)90016-1

Adkins, C. L., Russell, C. J., & Werbel, J. D. (1994). Judgments of fit in the selection process: The role of work value congruence. *Personnel Psychology*, *47*(3), 605–623. https://doi.org/10.1111/j.1744-6570.1994.tb01740.x

AERA, APA, & NCME (Eds.). (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.

Aken, J. E. van, Berends, H., & Bij, H. van der. (2010). *Problem-solving in organizations: A methodological handbook for business students* (4th ed.). Cambridge: Cambridge University Press.

Allison, G. T. (1969). Conceptual models and the cuban missile crisis. *American Political Science Review*, *63*(3), 689–718. https://doi.org/10.1017/S000305540025853X

Alvesson, M., & Sandberg, J. (2014). Habitat and habitus: Boxed-in versus box-breaking research. *Organization Studies*, *35*(7), 967–987. https://doi.org/10.1177/0170840614530916

Amelingmeyer, J. (2002). *Wissensmanagement: Analyse und Gestaltung der Wissensbasis von Unternehmen [Knowledge management: Analysis and design of the knowledge basis of companies]* (2nd ed.). Wiesbaden: Deutscher Universitätsverlag.

Amelung, D., & Funke, J. (2013). Dealing with the uncertainties of climate engineering: Warnings from a psychological complex problem solving perspective. *Technology in Society*, *35*(1), 32–40. https://doi.org/10.1016/j.techsoc.2013.03.001

Anakwe, U. P., Hall, J. C., & Schor, S. M. (2000). Knowledge-related skills and effective career management. *International Journal of Manpower*, *21*(7), 566–579. https://doi.org/10.1108/01437720010379024

Anderson, J. A. (2006). *Automata theory with modern applications*. Cambridge: Cambridge University Press.

Anderson, P. (1999). Complexity theory and organization science. *Organization Science*, *10*(3), 216–232. https://doi.org/10.1287/orsc.10.3.216

Annett, J., & Duncan, K. (1967). Task analysis and training design. *Occupational Psychology*, *41*, 211–221.

Arthur, M. B., & Rousseau, D. M. (1996). *The boundaryless career: A new employment principle for a new organizational era*. Oxford: Oxford University Press.

Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, *56*(1), 125–153. https://doi.org/10.1111/j.1744-6570.2003.tb00146.x

Ashworth, M. J. (2006). Preserving knowledge legacies: Workforce aging, turnover and human resource issues in the US electric power industry. *International Journal of Human Resource Management*, *17*(9), 1659–1688. https://doi.org/10.1080/09585190600878600

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *The American Economic Review*, *103*(5), 1553–1597. https://doi.org/10.1257/aer.103.5.1553

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the U.S. labor market. *American Economic Review*, *96*(2), 189–194. https://doi.org/10.1257/000282806777212620

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

Autor, D. H., & Price, B. (2013, June 21). *The changing task composition of the US labor market: An update of Autor, Levy and Murnane (2003)*. Unpublished manuscript, Cambridge. Retrieved from http://economics.mit.edu/files/9758

Axley, S. R., & McMahon, T. R. (2006). Complexity: A frontier for management education. *Journal of Management Education*, *30*(2), 295–315. https://doi.org/10.1177/1052562905277316

Badke-Schaub, P., Hofinger, G., & Lauche, K. (2012). *Human Factors: Psychologie sicheren Handelns in Risikobranchen [Human factors: Psychology of*. Heidelberg: Springer.

Baggen, Y., Mainert, J., Lans, T., Biemans, H. J. A., Greiff, S., & Mulder, M. (2015). Linking complex problem solving to opportunity identification competence within the context of entrepreneurship. *International Journal of Lifelong Education*, *34*(4), 412–429. https://doi.org/10.1080/02601370.2015.1060029

Baron, R. A. (2014). Thinking about cognition and its central role in entrepreneurship: Confessions of a “reformed” behaviorist. In J. R. Mitchell, R. K. Mitchell, & B. Randolph-Seng (Eds.), *Handbook of entrepreneurial cognition* (pp. 61–85). Cheltenham: Edward Elgar.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barth, C. M., & Funke, J. (2010). Negative affective environments improve complex solving performance. *Cognition & Emotion*, *24*(7), 1259–1268. https://doi.org/10.1080/02699930903223766

Bass, B. M. (1974). The substance and the shadow. *American Psychologist*, *29*(12), 870–886. https://doi.org/10.1037/h0037515

Baumgarten, D., Geishecker, I., & Görg, H. (2010). Offshoring, tasks, and the skill-wage pattern. *CEGE Discussion Paper*, *98*. Retrieved from http://webdoc.gwdg.de/ebook/serien/lm/CEGE/98.pdf

Becker, S. O., Ekholm, K., & Muendler, M.-A. (2013). Offshoring and the onshore composition of tasks and skills. *Journal of International Economics*, *90*(1), 91–106. https://doi.org/10.1016/j.jinteco.2012.10.005

Beckmann, J. F., & Goode, N. (2013). 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*, 1–20. https://doi.org/10.1007/s11251-013-9280-7

Bedeian, A. G., & Wren, D. A. (2002). Most influential management books of the 20th century. *Organizational Dynamics*, *29*(3), 221–225. https://doi.org/10.1016/S0090-2616(01)00022-5

Bird, A. (1994). Careers as repositories of knowledge: A new perspective on boundaryless careers. *Journal of Organizational Behavior*, *15*(4), 325–344. https://doi.org/10.1002/job.4030150404

Boreham, N. (2002). Work process knowledge, curriculum control and the work-based route to vocational qualifications. *British Journal of Educational Studies*, *50*(2), 225–237. https://doi.org/10.1111/1467-8527.00199

Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Job and work analysis: Methods, research, and applications for human resource management* (2nd ed., Vol. XII). Thousand Oaks: SAGE.

Braune, R., & Foshay, W. R. (1983). Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. *Instructional Science*, *12*(2), 121–145. https://doi.org/10.1007/BF00122453

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, *81*(3), 211–241. https://doi.org/10.1016/0001-6918(92)90019-A

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, *9*(2–3), 171–184. https://doi.org/10.1016/0747-5632(93)90005-D

Brockmann, M., Clarke, L., & Winch, C. (2008). Knowledge, skills, competence: European divergences in vocational education and training (VET)—the English, German and Dutch cases. *Oxford Review of Education*, *34*(5), 547–567. https://doi.org/10.1080/03054980701782098

Brouwers, S. A., & Van De Vijver, F. J. r. (2012). Intelligence 2.0 in I–O psychology: Revival or contextualization? *Industrial and Organizational Psychology*, *5*(2), 158–160. https://doi.org/10.1111/j.1754-9434.2012.01422.x

Brown, K. G., Le, H., & Schmidt, F. L. (2006). Specific aptitude theory revisited: Is there incremental validity for training performance? *International Journal of Selection and Assessment*, *14*(2), 87–100. https://doi.org/10.1111/j.1468-2389.2006.00336.x

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* (pp. 27–63). Hillsdale: Erlbaum.

Buchner, A., & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. *The Quarterly Journal of Experimental Psychology*, *46*(1), 83–118. https://doi.org/10.1080/14640749308401068

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, *13*(1), 40–52. https://doi.org/10.5465/AMR.1988.4306775

Campion, M. A., Fink, A. A., Ruggeberg, B. J., Carr, L., Phillips, G. M., & Odman, R. B. (2011). Doing competencies well: Best practices in competency modeling. *Personnel Psychology*, *64*(1), 225–262. https://doi.org/10.1111/j.1744-6570.2010.01207.x

Caplan, R. D. (1987). Person-environment fit theory and organizations: Commensurate dimensions, time perspectives, and mechanisms. *Journal of Vocational Behavior*, *31*(3), 248–267. https://doi.org/10.1016/0001-8791(87)90042-X

Cappelli, P. (2008). Talent management for the twenty-first century. *Harvard Business Review*, *86*(3), 74–81.

Cascio, W. F. (1995). Whither industrial and organizational psychology in a changing world of work? *American Psychologist*, *50*(11), 928–939. https://doi.org/10.1037/0003-066X.50.11.928

Cascio, W. F. (1998). *Managing human resources: Productivity, quality of work life, profits*. Boston: Irwin McGraw-Hill.

Chan, D. (1996). Cognitive misfit of problem-solving style at work: A facet of person-organization fit. *Organizational Behavior and Human Decision Processes*, *68*(3), 194–207. https://doi.org/10.1006/obhd.1996.0099

Chapelle, C. A., Enright, M. K., & Jamieson, J. (2010). Does an argument-based approach to validity make a difference? *Educational Measurement: Issues and Practice*, *29*(1), 3–13. https://doi.org/10.1111/j.1745-3992.2009.00165.x

Chatman, J. A. (1989). Matching people and organizations: Selection and socialization in public accounting firms. *Academy of Management Proceedings*, *1989*(1), 199–203. https://doi.org/10.5465/AMBPP.1989.4980837

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. https://doi.org/10.1111/1467-8624.00081

Churchman, C. W., Ackoff, R. L., & Arnoff, E. L. (1971). *Operations Research* (5th ed.). München: Oldenbourg.

Clifford, J. P. (1994). Job analysis: Why do it, and how should it be done? *Public Personnel Management*, *23*(2), 321–340. https://doi.org/10.1177/009102609402300211

Cohen, D. J. (2007). The very separate worlds of academic and practitioner publications in human resource management: Reasons for the divide and concrete solutions for bridging the gap. *Academy of Management Journal*, *50*(5), 1013–1019. https://doi.org/10.5465/AMJ.2007.27151946

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, *17*(1), 1–25. https://doi.org/10.2307/2392088

Collings, D. G., & Mellahi, K. (2009). Strategic talent management: A review and research agenda. *Human Resource Management Review*, *19*(4), 304–313. https://doi.org/10.1016/j.hrmr.2009.04.001

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653–665. https://doi.org/10.1016/0191-8869(92)90236-I

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. https://doi.org/10.1037/h0040957

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

De Fruyt, F., Wille, B., & John, O. P. (2015). Employability in the 21st century: Complex (interactive) problem solving and other essential skills. *Industrial and Organizational Psychology*, *8*(02), 276–281. https://doi.org/10.1017/iop.2015.33

De Meuse, K. P., Dai, G., & Hallenbeck, G. S. (2010). Learning agility: A construct whose time has come. *Consulting Psychology Journal: Practice and Research*, *62*(2), 119–130. https://doi.org/10.1037/a0019988

De Obeso Orendain, A. (2014). *Cognitive modelling of complex problem solving behaviour* (Dissertation). University of Sussex, Sussex.

Denison, D. R., Hooijberg, R., & Quinn, R. E. (1995). Paradox and performance: Toward a theory of behavioral complexity in managerial leadership. *Organization Science*, *6*(5), 524–540. https://doi.org/10.1287/orsc.6.5.524

DeRue, D. S., Ashford, S. J., & Myers, C. G. (2012). Learning agility: In search of conceptual clarity and theoretical grounding. *Industrial and Organizational Psychology*, *5*(3), 258–279. https://doi.org/10.1111/j.1754-9434.2012.01444.x

Dierdorff, E. C., & Wilson, M. A. (2003). A meta-analysis of job analysis reliability. *Journal of Applied Psychology*, *88*(4), 635–646. https://doi.org/10.1037/0021-9010.88.4.635

Dilchert, S., & Ones, D. S. (2009). Assessment center dimensions: Individual differences correlates and meta-analytic incremental validity. *International Journal of Selection and Assessment*, *17*(3), 254–270. https://doi.org/10.1111/j.1468-2389.2009.00468.x

Dörner, D. (1979a). *Problemlösen als Informationsverarbeitung [Problem solving as information processing]*. Stuttgart: Kohlhammer.

Dörner, D. (1979b). *Programm TAILORSHOP in der Version für TI-59 mit Drucker PC-100. Modifizierte und kommentierte Fassung von Norbert Streitz. [Program TAILORSHOP for TI-59 with printer PC-100. Modified and commented version by Norbert Streitz]*. Aachen: unpublished.

Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Games*, *11*(1), 87–106. https://doi.org/10.1177/104687818001100108

Dörner, D. (1989a). Die kleinen grünen Schildkröten und die Methoden der experimentellen Psychologie [Little green turtles und methods of experimental psychology]. *Sprache & Kognition*, *8*(2), 86–97.

Dörner, D. (1989b). *Die Logik des Misslingens [The logic of failure]*. Hamburg: Rowohlt.

Dörner, D. (1990). The logic of failure. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *327*(1241), 463–473. https://doi.org/10.1098/rstb.1990.0089

Dörner, D., & Kreuzig, H. W. (1983). Problemlösefähigkeit und Intelligenz. [Problem solving ability and intelligence.]. *Psychologische Rundschau*, *34*(4), 185–192.

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

Dörner, D., & Reither, F. (1978). Über das Problemlösen in sehr komplexen Realitätsbereichen. *Zeitschrift Für Experimentelle Und Angewandte Psychologie*, *25*, 527–551.

Dörner, D., & Wearing, A. J. (1995). Complex problem solving: Toward a (computer-simulated) theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The european perspective* (pp. 65–99). Hillsdale: Erlbaum.

Dries, N., Vantilborgh, T., & Pepermans, R. (2012). The role of learning agility and career variety in the identification and development of high potential employees. *Personnel Review*, *41*(3), 340–358. https://doi.org/10.1108/00483481211212977

Duijnhoven, H., & Neef, M. (2016). Disentangling wicked problems: A reflexive approach towards resilience governance. In A. J. Masys (Ed.), *Applications of systems thinking and soft operations research in managing complexity* (pp. 91–106). Heidelberg: Springer. https://doi.org/10.1007/978-3-319-21106-0_5

Dunbar, K. (1998). Problem solving. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 289–298). London: Blackwell.

Duncker, K. (1945). On problem-solving. *Psychological Monographs*, *58*(5), 1–113. https://doi.org/10.1037/h0093599

Dunnette, M. D. (1990). Blending the science and practice of industrial and organizational psychology: Where are we and where are we going? In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 1–27). Palo Alto: Consulting Psychologists Press.

Ebbesen, E. B., & Konečni, V. J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T. S. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (pp. 21–45). Hillsdale: Erlbaum.

Edwards, J. R. (1991). Person-job fit: A conceptual integration, literature review, and methodological critique. In C. L. Cooper & I. T. Robertson (Eds.), *International*

Review of Industrial/Organizational Psychology (Vol. 9, pp. 383–357). London: Wiley.

Eichinger, R. W., & Lombardo, M. M. (2004). Learning agility as a prime indicator of potential. *Human Resource Planning*, *27*(4), 12–15.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods*, *8*(1), 38–60. https://doi.org/10.1037/1082-989X.8.1.38

Elio, R., & Scharf, P. B. (1990). Modeling novice-to-expert shifts in problem-solving strategy and knowledge organization. *Cognitive Science*, *14*(4), 579–639. https://doi.org/10.1016/0364-0213(90)90010-T

Elqayam, S., & Evans, J. S. B. T. (2011). Subtracting “ought” from “is”: Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, *34*(05), 233–248. https://doi.org/10.1017/S0140525X1100001X

Elster, J. (1987). *Subversion der Rationalität [Subversion of rationality]*. Frankfurt am Main: Campus.

Engelhart, M., Funke, J., & Sager, S. (2013). A decomposition approach for a new test-scenario in complex problem solving. *Journal of Computational Science*, *4*, 245–254. https://doi.org/10.1016/j.jocs.2012.06.005

Ernst, G. W., & Newell, A. (1969). *GPS: A case study in generality and problem solving*. New York: Academic Press.

Fisch, R., & Beck, D. (2004). *Komplexitätsmanagement*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Fischer, A. (2015). *Assessment of problem solving skills by means of multiple complex systems – Validity of finite automata and linear dynamic systems* (Dissertation). University of Heidelberg, Heidelberg.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, *4*(1), 19–42. https://doi.org/10.7771/1932-6246.1118

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. https://doi.org/10.1016/j.lindif.2015.02.008

Fischer, A., Holt, D., & Funke, J. (2015). Promoting the growing field of dynamic decision making. *Journal of Dynamic Decision Making*, *1*(1), 1–3. https://doi.org/10.11588/jddm.2015.1.23807

Fischer, A., & Neubert, J. C. (2015). The multiple faces of complex problems: A model of complex problem solving competency and its assessment. *Journal of Dynamic Decision Making*, *1*(6), 1–14. https://doi.org/10.11588/jddm.2015.1.23945

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, *81*(1), 59–74. https://doi.org/10.1037/h0035872

Fleishman, E. A., & Reilly, M. E. (1992). *Handbook of human abilities: Definitions, measurements, and job task requirements.* Palo Alto: Consulting Psychologists Press.

Foss, N. J., & Weber, L. (2016). Moving opportunism to the back seat: Bounded rationality, costly conflict, and hierarchical forms. *Academy of Management Review*, *41*(1), 61–79. https://doi.org/10.5465/amr.2014.0105

Frensch, P. A., & Funke, J. (1995a). *Complex problem solving: The european perspective*. Hillsdale: Erlbaum.

Frensch, P. A., & Funke, J. (1995b). 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: Erlbaum.

Frischkorn, G., Greiff, S., & Wüstenberg, S. (2014). The development of complex problem solving: A latent growth curve analysis. *Journal of Educational Psychology*, *106*, 1007–1020. https://doi.org/doi:10.1037/a0037114

Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? [Some remarks on the problems of problem solving research, or: Is intelligence a predictor?]. *Diagnostica*, *29*, 283–302.

Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle [Control of dynamic systems through subjective causal models]. *Zeitschrift Für Psychologie*, *193*(4), 435–457.

Funke, J. (1992). *Wissen über dynamische Systeme. Erwerb, Repräsentation und Anwendung [Knowledge about dynamic systems: Acquisition, representation, and use]*. Heidelberg: Springer.

Funke, J. (1993). Computergestützte Arbeitsproben: Begriffsklärung, Beispiele sowie Entwicklungspotentiale [Computer-based work samples: Definition, examples, and future trends]. *Zeitschrift Für Arbeits- Und Organisationspsychologie*, *37*, 119–129.

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. (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.

Funke, J. (2008). Suchprozesse beim Problemlösen: Dem Denken auf der Spur [Search processes during problem solving: Following the trail of thinking]. In G. Jüttemann (Ed.), *Suchprozesse der Seele. Die Psychologie des Erwägens* (pp. 71–81). Göttingen: Vandenhoeck & Ruprecht.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

Funke, J. (2014a). Analysis of minimal complex systems and complex problem solving require different forms of causal cognition. *Cognitive Science*, *5*(739), 1–3. https://doi.org/10.3389/fpsyg.2014.00739

Funke, J. (2014b). Problem solving: What are the important questions? In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 493–498). Austin: Cognitive Science Society.

Funke, J., & Buchner, A. (1992). Finite Automaten als Instrumente für die Analyse von wissensgeleiteten Problemlöseprozessen: Vorstellung eines neuen Untersuchungsparadigmas [Finite-state automata as instruments for the analysis of knowledge based problem solving: Introducing a new research paradigm]. *Sprache & Kognition*, *11*(1), 27–37.

Funke, J., Fischer, A., & Holt, D. V. (in preparation). Competencies for complexity: Problem solving in the 21st century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and Teaching of 21st Century Skills* (Vol. 3). New York: Springer.

Funke, U. (1993). Computergestützte Eignungsdiagnostik mit komplexen dynamischen Szenarios [Computer-based personnel selection with complex dynamic scenarios]. *Zeitschrift Für Arbeits- Und Organisationspsychologie*, *37*(3), 109–118.

Funke, U. (1995). Using complex problem solving tasks in personnel selection and training. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The european perspective* (pp. 219–240). Hillsdale: Erlbaum.

Gaschler, R. (2009). *Information reduction as item-general strategy change* (Dissertation). Humboldt-Universität zu Berlin, Berlin.

Gathmann, C., & Schönberg, U. (2010). How general is human capital? A task-based approach. *Journal of Labor Economics*, *28*(1), 1–49. https://doi.org/10.1086/649786

Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712. https://doi.org/10.1002/smj.475

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Research Policy*, *31*(8–9), 1257–1274. https://doi.org/10.1016/S0048-7333(02)00062-8

Geels, F. W. (2010). Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective. *Research Policy*, *39*(4), 495–510. https://doi.org/10.1016/j.respol.2010.01.022

Geels, F. W., & Schot, J. (2007). Typology of sociotechnical transition pathways. *Research Policy*, *36*(3), 399–417. https://doi.org/10.1016/j.respol.2007.01.003

Gell-Mann, M. (1995). What is complexity? Remarks on simplicity and complexity. *Complexity*, *1*(1), 16–19. https://doi.org/10.1002/cplx.6130010105

Gick, M. L., & Holyoak, K. . (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*(1), 451–482. https://doi.org/10.1146/annurev-psych-120709-145346

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*(4), 591–635. https://doi.org/10.1016/S0364-0213(03)00031-4

Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, *21*(2), 273–286. https://doi.org/10.1016/j.chb.2004.02.014

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *Review of Economics and Statistics*, *89*(1), 118–133. https://doi.org/10.1162/rest.89.1.118

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *The American Economic Review*, *99*(2), 58–63. https://doi.org/10.1257/aer.99.2.58

Gottfredson, L. S. (1997a). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23. https://doi.org/10.1016/S0160-2896(97)90011-8

Gottfredson, L. S. (1997b). Why g matters: The complexity of everyday life. *Intelligence*, *24*(1), 79–132. https://doi.org/10.1016/S0160-2896(97)90014-3

Gray, W. D. (2000). The nature and processing of errors in interactive behavior. *Cognitive Science*, *24*(2), 205–248. https://doi.org/10.1207/s15516709cog2402_2

Greenhaus, J. H., Callanan, G. A., & Kaplan, E. (1995). The role of goal setting in career management. *International Journal of Career Management*, *7*(5), 3–12. https://doi.org/10.1108/09556219510093285

Greiff, S. (2012). *Individualdiagnostik der komplexen Problemlösefähigkeit [Diagnostics of problem solving ability on an individual level]*. Münster: Waxmann.

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., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence*, *41*(5), 579–596. https://doi.org/10.1016/j.intell.2013.07.012

Greiff, S., & Funke, J. (2009). Measuring complex problem solving: the MicroDYN approach. In F. Scheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing* (pp. 157–163). Luxembourg: Office for Official Publications of the European Communities.

Greiff, S., & Funke, J. (2010). Systematische Erforschung komplexer Problemlösefähigkeit anhand minimal komplexer Systeme [Systematic research on complex problem solving by means of minimal complex systems]. In E. Klieme, D. Leutner, & M. Kenk (Eds.), *Kompetenzmodellierung* (pp. 216–227). Weinheim: Beltz.

Greiff, S., Holt, D. V., & Funke, J. (2013). Perspectives on problem solving in educational assessment: Analytical, interactive, and collaborative problem solving. *The Journal of Problem Solving*, *5*(2), 71–91. https://doi.org/10.7771/1932-6246.1153

Greiff, S., Kretzschmar, A., Müller, J. C., Spinath, B., & Martin, R. (2014). Computer-based assessment of Complex Problem Solving in educational contexts and how it is influenced by students' level of Information and Communication Technology literacy. *Journal of Educational Psychology*, *106*(3), 666–680. https://doi.org/10.1037/a0035426

Greiff, S., & Martin, R. (2014). What you see is what you (don't) get: A comment on Funke's (2014) opinion paper. *Cognitive Science*, *5*(1120), 1–3. https://doi.org/10.3389/fpsyg.2014.01120

Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, *36*, 37–48. https://doi.org/10.1016/j.lindif.2014.08.003

Greiff, S., Neubert, J. C., Niepel, C., & Ederer, P. (2015). Problem solving - facilitating the utilization of a concept towards lifelong education. *International Journal of Lifelong Education*, *34*(4), 373–375. https://doi.org/10.1080/02601370.2015.1060028

Greiff, S., & Wüstenberg, S. (2015). Assessment with microworlds using MicroDYN: Measurement invariance and latent mean comparisons: psychometric properties across several student samples and blue-collar workers. *European Journal of Psychological Assessment*, *30*, 304–314. https://doi.org/10.1027/1015-5759/a000194

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

Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21st century skills*. New York: Springer.

Grossman, G. M., & Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, *98*(5), 1978–1997. https://doi.org/10.1257/aer.98.5.1978

Guarana, C. L., & Hernandez, M. (2015). Building sense out of situational complexity The role of ambivalence in creating functional leadership processes. *Organizational Psychology Review*, *5*(1), 50–73. https://doi.org/10.1177/2041386614543345

Güss, C. D., Tuason, M. T., & Orduña, L. V. (2015). Strategies, tactics, and errors in dynamic decision making. *Journal of Dynamic Decision Making*, *1*(3), 1–14.

Gustafsson, V. (2006). *Entrepreneurial decision-making: Individuals, tasks and cognitions*. Cheltenham: Elgar.

Hacker, W., & von der Weth, R. (2012). Denken – Entscheiden – Handeln [Thinking - deciding - acting]. In P. Badke-Schaub, G. Hofinger, & K. Lauche (Eds.), *Human Factors* (pp. 83–99). Heidelberg: Springer.

Hackman, J. R. (1969). Toward understanding the role of tasks in behavioral research. *Acta Psychologica*, *31*, 97–128. https://doi.org/10.1016/0001-6918(69)90073-0

Hærem, T., Pentland, B. T., & Miller, K. D. (2015). Task complexity: Extending a core concept. *Academy of Management Review*, *40*(3), 446–460. https://doi.org/10.5465/amr.2013.0350

Hafenbrädl, S., Waeger, D., Marewski, J. N., & Gigerenzer, G. (2016). Applied decision making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition*, *5*(2), 215–231. https://doi.org/10.1016/j.jarmac.2016.04.011

Hall, D. T., & Mirvis, P. H. (1996). The new protean career. In D. T. Hall (Ed.), *The Career is Dead - Long Live the Career* (pp. 15–45). San Francisco: Jossey-Bass.

Harvey, D. F., & Bowin, R. B. (1996). *Human resource management: An experiential approach*. Englewood Cliffs: Prentice-Hall.

Hasselmann, D. (1993). *Computersimulierte komplexe Problemstellungen in der Management-Diagnostik: Die Theorie zum Konzept Herausforderung Komplexität [Computer simulated complex problems for management assessment: The theory for the challenge of complexity]*. Hamburg: Windmühle.

Heller, K. A., Kratzmeier, H., & Lengfelder, A. (1998). *Matrizen-Test-Manual. Ein Handbuch zu den Standard Progressive Matrices von J.C. Raven [Matrices test manual. A handbook of Standard Progressive Matrices of J.C. Raven]*. Göttingen: Beltz.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. [Cognitive Abilities Test (CogAT; Thorndike, L. & Hagen, E., 1954-1986) - German adapted version]*. Göttingen: Beltz.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 37–56). Dordrecht: Springer. https://doi.org/10.1007/978-94-017-9395-7_2

Hillier, F. S., & Lieberman, G. J. (2005). *Introduction to operations research* (8th ed.). Boston: McGraw-Hill.

Holyoak, K. J. (2005). Analogy. In K. . Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). New York: Cambridge University Press.

Hossiep, R., & Schulte, M. (2008). *BOWIT. Bochumer Wissenstest [BOWIT. Bochum knowledge test].* Göttingen: Hogrefe.

Howard, A. (Ed.). (1995). *The changing nature of work*. San Francisco: Jossey-Bass.

Hundertmark, J., Holt, D. V., Fischer, A., Said, N., & Fischer, H. (2015). System structure and cognitive ability as predictors of performance in dynamic system control tasks. *Journal of Dynamic Decision Making*, *1*(5), 1–10. https://doi.org/10.11588/jddm.2015.1.26416

Hunt, E., & Madhyastha, T. M. (2012). Cognitive demands of the workplace. *Journal of Neuroscience, Psychology, and Economics*, *5*(1), 18–37. https://doi.org/10.1037/a0026177

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior*, *29*(3), 340–362. https://doi.org/10.1016/0001-8791(86)90013-8

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, *2*(3–4), 447–472. https://doi.org/10.1037/1076-8971.2.3-4.447

Hutchins, E. (1996). *Cognition in the wild*. Cambridge: MIT Press.

Ikenaga, T., & Kambayashi, R. (2010). Long-term trends in the polarization of the japanese labor market: The increase of non-routine task input and its valuation in the labor market. *Hitotsubashi University Institute of Economic Research Working Paper*. Retrieved from http://takayama-online.net/pie/stage3/Japanese/d_p/dp2009/dp464/text2.pdf

Isenberg, D. J. (1986). Thinking and managing: A verbal protocol analysis of managerial problem solving. *The Academy of Management Journal*, *29*(4), 775–788. https://doi.org/10.2307/255944

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

Jenkins, D. P. (2009). *Cognitive work analysis coping with complexity*. Farnham: Ashgate.

Johns, G. (2001). In praise of context. *Journal of Organizational Behavior*, *22*(1), 31–42. https://doi.org/10.1002/job.80

Jonassen, D. H. (2004). *Learning to solve problems: An instructional design guide*. San Francisco: Wiley.

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, *87*(4), 765–780. https://doi.org/10.1037/0021-9010.87.4.765

Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, *89*(3), 542–552. https://doi.org/10.1037/0021-9010.89.3.542

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3), 621–652. https://doi.org/10.1111/j.1744-6570.1999.tb00174.x

Kallerud, E., Amanatidou, E., Upham, P., Nieminen, M., Klitkou, A., Olsen, D. S., … Scordato, L. (2013). Dimensions of research and innovation policies to address grand and global challenges. *NIFU, Oslo*. Retrieved from http://www.euspri-forum.eu/key_missions/CPRI_Position_paper.pdf

Kämmer, J. E., Gaissmaier, W., & Czienskowski, U. (2013). The environment matters: Comparing individuals and dyads in their adaptive use of decision strategies. *Judgment and Decision Making*, *8*(3), 299–329.

Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, *38*(4), 319–342. https://doi.org/10.1111/j.1745-3984.2001.tb01130.x

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport: Praeger.

Kaslow, N. J., Rubin, N. J., Bebeau, M. J., Leigh, I. W., Lichtenberg, J. W., Nelson, P. D., … Leon, I. (2007). Guiding principles and recommendations for the assessment of competence. *Professional Psychology: Research and Practice*, *38*(5), 441–451. https://doi.org/10.1037/0735-7028.38.5.441

Keegan, J., & Kabanoff, B. (2007). Indirect industry-and subindustry-level managerial discretion measurement. *Organizational Research Methods*, *11*(4), 682–694. https://doi.org/10.1177/1094428107308897

Kersting, M. (1999). *Diagnostik und Personalauswahl mit computergestützten Problemlöseszenarien? [Diagnosis and personnel selection by means of computer-based problem-solving scenarios]*. Göttingen: Hogrefe.

Kersting, M. (2001). Zur Konstrukt- und Kriteriumsvalidität von Problemlöseszenarien anhand der Vorhersage von Vorgesetztenurteilen über die berufliche Bewährung [On the construct and criterion validity of problem-solving scenarios based on the prediction of supervisor assessment of job performance]. *Diagnostica*, *47*(2), 67–76. https://doi.org/10.1026//0012-1924.47.2.67

Kersting, M., & Süß, H.-M. (1995). Kontentvalide Wissensdiagnostik und Problemlösen: Zur Entwicklung, testtheoretischen Begründung und empirischen Bewährung eines problemspezifischen Diagnoseverfahrens [Content-valid diagnosis of knowledge and problem solving: Development, test theoretical justification and empirical validation of a new problem-specific test]. *Zeitschrift Für Pädagogische Psychologie*, *9*(2), 83–94.

Kieser, A., Nicolai, A., & Seidl, D. (2015). The practical relevance of management research: Turning the debate on relevance into a rigorous scientific research program. *The Academy of Management Annals*, *9*(1), 143–233. https://doi.org/10.1080/19416520.2015.1011853

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, *78*(1), 85–123. https://doi.org/10.3102/0034654307313402

Klein, G. (2008). Naturalistic decision making. *Human Factors*, *50*(3), 456–460. https://doi.org/10.1518/001872008X288385

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: Models and methods*. Norwood: Ablex.

Klein, K. J., Dansereau, F., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review*, *19*(2), 195. https://doi.org/10.2307/258703

Klein, R., & Scholl, A. (2004). *Planung und Entscheidung: Konzepte, Modelle und Methoden einer modernen betriebswirtschaftlichen Entscheidungsanalyse. [Planning and decision making]*. München: Vahlen.

Klieme, E., Funke, J., Leutner, D., Reimann, P., & Wirth, J. (2001). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie [Problem-solving competence as a cross-curricula key competence]. *Zeitschrift Für Pädagogik*, *47*(2), 179–200.

Kluge, A. (2008a). Performance assessments with microworlds and their difficulty. *Applied Psychological Measurement*, *32*(2), 156–180. https://doi.org/10.1177/0146621607300015

Kluge, A. (2008b). 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

Kluwe, R. H., Misiak, C., & Haider, H. (1989). Modelling the process of complex system control. In P. M. Milling & E. O. K. Zahn (Eds.), *Computer-Based Management of Complex Systems* (pp. 335–342). Heidelberg: Springer.

Kohn, M. ., & Schooler, C. (1982). Job conditions and personality: A longitudinal assessment of their reciprocal effects. *American Journal of Sociology*, *87*(6), 1257–1286. https://doi.org/10.1086/227593

Kohn, M. L., & Schooler, C. (1978). The reciprocal effects of the substantive complexity of work and intellectual flexibility: A longitudinal assessment. *American Journal of Sociology*, *84*(1), 24–52. https://doi.org/10.2307/2777977

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*(2), 248–294. https://doi.org/10.1016/0010-0285(85)90009-X

Kretzschmar, A. (2015). *Konstruktvalidität des Komplexen Problemlösens unter besonderer Berücksichtigung moderner diagnostischer Ansätze [Construct validity of complex problem solving with a particular focus on modern assessment approaches]* (Dissertation). University of Luxembourg, Luxembourg.

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. https://doi.org/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*(1), 55–69. https://doi.org/10.1016/j.intell.2015.11.004

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, *49*(1), 1–49. https://doi.org/10.1111/j.1744-6570.1996.tb01790.x

Kröner, S., Plass, J., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, *33*(4), 347–368. https://doi.org/10.1016/j.intell.2005.03.002

Kuhn, D., Garcia-Mila, M., Zohar, A., Andersen, C., White, S. H., Klahr, D., & Carver, S. M. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, *60*(4), 1–157. https://doi.org/10.2307/1166059

Lang, J. W. B., Kersting, M., Hülsheger, U. R., & Lang, J. (2010). General mental ability, narrower cognitive abilities, and job performance: The perspective of the nested-factors model of cognitive abilities. *Personnel Psychology*, *63*(3), 595–640. https://doi.org/10.1111/j.1744-6570.2010.01182.x

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, *208*(4450), 1335–1342. https://doi.org/10.1126/science.208.4450.1335

Law, J., & Callon, M. (2006). Leben und Sterben eines Flugzeugs: ein Netzwerkanalyse technischen Wandels [The life and death of an aircraft: A network analysis of technical change]. In A. Belliger & D. J. Krieger (Eds.), *ANThology: ein einführendes Handbuch zur Akteur-Netzwerk-Theorie* (pp. 447–482). Berlin: transcript.

Leutner, D., Fleischer, J., Wirth, J., Greiff, S., & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien [Analytical and dynamic problem-solving competence from the perspective of studies on academic achievement in international comparison]. *Psychologische Rundschau*, *63*(1), 34–42. https://doi.org/10.1026/0033-3042/a000108

Leutner, D., Klieme, E., Meyer, K., & Wirth, J. (2004). Problemlösen [Problem solving]. In M. Prenzel & PISA-Konsortium Deutschland (Eds.), *PISA 2003: der Bildungsstand der Jugendlichen in Deutschland* (pp. 147–175). Münster: Waxmann.

Levine, E. L., Ash, R. A., Hall, H., & Sistrunk, F. (1983). Evaluation of job analysis methods by experienced job analysts. *The Academy of Management Journal*, *26*(2), 339–348. https://doi.org/10.2307/255981

Levine, E. L., Sistrunk, F., McNutt, K. J., & Gael, S. (1988). Exemplary job analysis systems in selected organizations: A description of process and outcomes. *Journal of Business and Psychology*, *3*(1), 3–21. https://doi.org/10.1007/BF01016745

Levine, E. L., Spector, P. E., Menon, S., Narayanan, L., & Cannon-Bowers, J. A. (1996). Validity generalization for cognitive, psychomotor, and perceptual tests for craft jobs in the utility industry. *Human Performance*, *9*(1), 1–22. https://doi.org/10.1207/s15327043hup0901_1

Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education*, *56*(3), 556–573. https://doi.org/10.1016/j.compedu.2010.09.015

Lewis, R. E., & Heckman, R. J. (2006). Talent management: A critical review. *Human Resource Management Review*, *16*(2), 139–154. https://doi.org/10.1016/j.hrmr.2006.03.001

Lipshitz, R., Klein, G., Orasanu, J., & Salas, E. (2001). Taking stock of naturalistic decision making. *Journal of Behavioral Decision Making*, *14*(5), 331–352. https://doi.org/10.1002/bdm.381

Lipshitz, R., & Strauss, O. (1997). Coping with uncertainty: A naturalistic decision-making analysis. *Organizational Behavior and Human Decision Processes*, *69*(2), 149–163. https://doi.org/10.1006/obhd.1997.2679

Liu, P., & Li, Z. (2012). Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics*, *42*(6), 553–568. https://doi.org/10.1016/j.ergon.2012.09.001

Loevinger, J. (1954). The attenuation paradox in test theory. *Psychological Bulletin*, *51*(5), 493–504. https://doi.org/10.1037/h0058543

Lombardo, M. M., & Eichinger, R. W. (2000). High potentials as high learners. *Human Resource Management*, *39*(4), 321–329. https://doi.org/10.1002/1099-050X(200024)39:4<321::AID-HRM4>3.0.CO;2-1

Lord, R. G., de Vader, C. L., & Alliger, G. M. (1986). A meta-analysis of the relation between personality traits and leadership perceptions: An application of validity generalization procedures. *Journal of Applied Psychology*, *71*(3), 402–410. https://doi.org/10.1037/0021-9010.71.3.402

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

Mainzer, K. (2009). Challenges of complexity in the 21st century. An interdisciplinary introduction. *European Review*, *17*(02), 219. https://doi.org/10.1017/S1062798709000714

March, J. G., & Simon, H. A. (1993). *Organizations* (2nd ed.). Cambridge: Blackwell.

Mayer, R. E. (2011). Problem solving and reasoning. In V. G. Aukrust (Ed.), *Learning and cognition in education* (pp. 112–117). Oxford: Elsevier.

McClelland, D. C. (1973). Testing for competence rather than for “intelligence.” *American Psychologist*, *28*(1), 1–14. https://doi.org/10.1037/h0034092

McCormick, E., Jeanneret, P., & Mecham, R. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*, *56*(4), 347–368. https://doi.org/10.1037/h0033099

McDermott, D. (1976). Artificial intelligence meets natural stupidity. *ACM SIGART Bulletin*, (57), 4–9. https://doi.org/10.1145/1045339.1045340

McElhaney, K. W., & Linn, M. C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching*, *48*(7), 745–770. https://doi.org/10.1002/tea.20423

McEntire, L. E., Dailey, L. R., Osburn, H. K., & Mumford, M. D. (2006). Innovations in job analysis: Development and application of metrics to analyze job data. *Human Resource Management Review*, *16*(3), 310–323. https://doi.org/10.1016/j.hrmr.2006.05.004

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004

Metcalf, L., & Benn, S. (2012). Leadership for sustainability: An evolution of leadership ability. *Journal of Business Ethics*, *112*(3), 369–384. https://doi.org/10.1007/s10551-012-1278-6

Mezias, J. M., & Starbuck, W. H. (2003). Studying the accuracy of managers’ perceptions: A research odyssey. *British Journal of Management*, *14*(1), 3–17. https://doi.org/10.1111/1467-8551.00259

Middleton, H. (2002). Complex problem solving in a workplace setting. *International Journal of Educational Research*, *37*(1), 67–84. https://doi.org/10.1016/S0883-0355(02)00022-8

Mintzberg, H. (1973). *The nature of managerial work*. New York: Harpercollins.

Mintzberg, H. (1975). The manager's job: Folklore and fact. *Harvard Business Review*, *53*(4), 49–61.

Morelli, N., Illingworth, A. J., & Handler, C. (2015). Questions about I-O Psychology's future. *Industrial and Organizational Psychology*, *8*(02), 269–276. https://doi.org/10.1017/iop.2015.32

Motowildo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, *10*(2), 71–83. https://doi.org/10.1207/s15327043hup1002_1

Müller, J. C., Kretzschmar, A., & Greiff, S. (2013). Exploring exploration: Inquiries into exploration behavior in complex problem solving assessment. In *Proceedings of the 6th International Conference on Educational Data Mining* (pp. 336–337). Memphis: International Educational Data Mining Society.

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*(2), 91–122. https://doi.org/10.1080/10400419109534380

Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. (2000). Leadership skills for a changing world: Solving complex social problems. *The Leadership Quarterly*, *11*(1), 11–35. https://doi.org/10.1016/S1048-9843(99)00041-7

Murphy, K. R., & Saal, F. E. (1990). What should we expect from scientist–practitioners? In K. R. Murphy & F. E. Saal (Eds.), *Psychology in organizations: Integrating science and practice* (pp. 49–66). Hillsdale, NJ: Erlbaum.

Naikar, N., Moylan, A., & Pearce, B. (2006). Analysing activity in complex systems with cognitive work analysis: Concepts, guidelines and case study for control task analysis. *Theoretical Issues in Ergonomics Science*, *7*(4), 371–394. https://doi.org/10.1080/14639220500098821

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. (Committee on Defining Deeper Learning and 21st Century Skills, J. W. Pellegrino, & M. L. Hilton, Eds.). Washington, DC: The National Academies Press.

Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2015). Extending the assessment of complex problem solving to finite state automata: Embracing heterogeneity. *European Journal of Psychological Assessment*, *31*(3), 181–194. https://doi.org/10.1027/1015-5759/a000224

Neubert, J. C., Lans, T., Mustafic, M., Greiff, S., & Ederer, P. (2017). Complex problem-solving in a changing world: Bridging domain-specific and transversal competence demands in vocational education. In M. Mulder (Ed.), *Competence-based Vocational and Professional Education* (pp. 953–969). Cham: Springer. https://doi.org/10.1007/978-3-319-41713-4_44

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*(2), 238–268. https://doi.org/10.1017/iop.2015.14

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Oxford: Prentice-Hall.

Nickerson, J. A., & Zenger, T. R. (2004). A knowledge-based theory of the firm: The problem-solving perspective. *Organization Science*, *15*(6), 617–632. https://doi.org/10.1287/orsc.1040.0093

Nilsson, N. J. (1971). *Problem-solving methods in artificial intelligence*. New York: McGraw-Hill.

Norman, D. A. (2013). *The design of everyday things*. New York: Basic Books.

Novick, L. R., & Bassok, M. (2005). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 321–349). New York: Cambridge University Press.

OECD. (2013a). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD Publishing.

OECD. (2013b). Problem-solving framework. In OECD, *PISA 2012 assessment and analytical framework* (pp. 119–137). Paris: OECD Publishing.

OECD. (2014). *PISA 2012 results: Creative problem solving* (Vol. V). Paris: OECD Publishing.

Oh, I.-S. (2015). Complex and collaborative problem solving: New, unique, and useful? *Industrial and Organizational Psychology*, *8*(02), 301–304. https://doi.org/10.1017/iop.2015.38

Ohlsson, S. (2012). The problems with problem solving: Reflections on the rise, current status, and possible future of a cognitive research paradigm. *The Journal of Problem Solving*, *5*(1), 101–128. https://doi.org/10.7771/1932-6246.1144

Öllinger, M., Hammon, S., Grundherr, M. von, & Funke, J. (2015). Does visualization enhance complex problem solving? The effect of causal mapping on performance in the computer-based microworld Tailorshop. *Educational Technology Research and Development*, 1–17. https://doi.org/10.1007/s11423-015-9393-6

O'Reilly, C. A., Chatman, J. A., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, *34*(3), 487–516. https://doi.org/10.2307/256404

Osterloh, M., & Frey, B. S. (2000). Motivation, knowledge transfer, and organizational forms. *Organization Science*, *11*(5), 538–550. https://doi.org/10.1287/orsc.11.5.538.15204

Oswald, F. L., & Hough, L. (2012). I–O 2.0 from intelligence 1.5: Staying (just) behind the cutting edge of intelligence theories. *Industrial and Organizational Psychology*, *5*(2), 172–175. https://doi.org/10.1111/j.1754-9434.2012.01425.x

Pearlman, K. (1980). Job families: A review and discussion of their implications for personnel selection. *Psychological Bulletin*, *87*(1), 1–28. https://doi.org/10.1037/0033-2909.87.1.1

Pentland, B. T., & Hærem, T. (2015). Organizational routines as patterns of action: Implications for organizational behavior. *Annual Review of Organizational*

Psychology and Organizational Behavior, *2*(1), 465–487. https://doi.org/10.1146/annurev-orgpsych-032414-111412

Pentland, B. T., Hærem, T., & Hillison, D. (2011). The (n)ever-changing world: Stability and change in organizational routines. *Organization Science*, *22*(6), 1369–1383. https://doi.org/10.1287/orsc.1110.0624

Peterson, H. C. (2009). Transformational supply chains and the “wicked problem” of sustainability: aligning knowledge, innovation, entrepreneurship, and leadership. *Journal on Chain and Network Science*, *9*(2), 71–82. https://doi.org/10.3920/JCNS2009.x178

Peterson, N. G., & American Psychological Association. (1999). *An occupational information system for the 21st century: The development of O*NET*. Washington, D.C.: American Psychological Association.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. https://doi.org/10.1037/a0014996

Probst, G., Raub, S., & Romhardt, K. (2006). *Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen [Managing knowledge]* (5., überarb. Aufl). Wiesbaden: Gabler.

Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg. [The relation between test intelligence and problem solving success.]. *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*, *189*(1), 79–100.

Quesada, J., Kintsch, W., & Gomez, E. (2005). Complex problem-solving: A field in search of a definition? *Theoretical Issues in Ergonomics Science*, *6*(1), 5–33. https://doi.org/10.1080/14639220512331311553

Raffiee, J., & Coff, R. (2016). Micro-foundations of firm-specific human capital: When do employees perceive their skills to be firm-specific? *Academy of Management Journal*, *59*(3), 766–790. https://doi.org/10.5465/amj.2014.0286

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science*, *27*(2–3), 183–213. https://doi.org/10.1016/S0925-7535(97)00052-0

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. New York: Wiley.

Rausch, A., Seifried, J., Wuttke, E., Kögler, K., & Brandt, S. (2016). Reliability and validity of a computer-based assessment of cognitive and non-cognitive facets of problem-solving competence in the business domain. *Empirical Research in Vocational Education and Training*, *8*(1), 9. https://doi.org/10.1186/s40461-016-0035-y

Raven, J. (2000). Psychometrics, cognitive ability, and occupational performance. *Review of Psychology*, *7*(1–2), 51–74.

Ready, D. A., & Conger, J. A. (2007). Make your company a talent factory. *Harvard Business Review*, *85*(6), 68–77.

Ree, M. J., & Carretta, T. R. (2002). g2K. *Human Performance*, *15*(1–2), 3–23. https://doi.org/10.1080/08959285.2002.9668081

Reed, S. K. (2012). Learning by mapping across situations. *Journal of the Learning Sciences*, *21*(3), 353–398. https://doi.org/10.1080/10508406.2011.607007

Reed, S. K., Ernst, G. W., & Banerji, R. (1974). The role of analogy in transfer between similar problem states. *Cognitive Psychology*, *6*(3), 436–450. https://doi.org/10.1016/0010-0285(74)90020-6

Reeve, C. L. (2004). Differential ability antecedents of general and specific dimensions of declarative knowledge: More than g. *Intelligence*, *32*(6), 621–652. https://doi.org/10.1016/j.intell.2004.07.006

Reiter-Palmon, R., & Illies, J. J. (2004). Leadership and creativity: Understanding leadership from a creative problem-solving perspective. *The Leadership Quarterly*, *15*(1), 55–77. https://doi.org/10.1016/j.leaqua.2003.12.005

Rich, E. (2008). *Automata, computability and complexity: Theory and applications*. Upper Saddle River: Prentice Hall.

Rigas, G., Carling, E., & Brehmer, B. (2002). Reliability and validity of performance measures in microworlds. *Intelligence*, *30*(5), 463–480. https://doi.org/10.1016/S0160-2896(02)00121-6

Riggio, R. E., & Saggi, K. (2015). Incorporating “soft skills” into the collaborative problem-solving equation. *Industrial and Organizational Psychology*, *8*(02), 281–284. https://doi.org/10.1017/iop.2015.34

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/BF01405730

Robbins, P., & Aydede, M. (2008). *Cambridge handbook of situated cognition*. Cambridge: Cambridge University Press.

Rodriguez, D., Patel, R., Bright, A., Gregory, D., & Gowing, M. K. (2002). Developing competency models to promote integrated human resource practices. *Human Resource Management*, *41*(3), 309–324. https://doi.org/10.1002/hrm.10043

Rollett, W. (2008). *Strategieeinsatz, erzeugte Information und Informationsnutzung bei der Exploration und Steuerung komplexer dynamischer Systeme [Use of strategy, generated information and use of information when exploring and controlling complex dynamic systems]*. Münster: LIT.

Roth, W.-M. (1998). Situated cognition and assessment of competence in science. *Evaluation and Program Planning*, *21*(2), 155–169. https://doi.org/10.1016/S0149-7189(98)00004-4

Roth, W.-M. (2001). Situating cognition. *Journal of the Learning Sciences*, *10*(1–2), 27–61. https://doi.org/10.1207/S15327809JLS10-1-2_4

Rouwette, E. A. J. A., Größler, A., & Vennix, J. A. M. (2004). Exploring influencing factors on rationality: A literature review of dynamic decision-making studies in system dynamics. *Systems Research and Behavioral Science*, *21*(4), 351–370. https://doi.org/10.1002/sres.647

Rouwette, E. A. J. A., & Vennix, J. A. M. (2006). System dynamics and organizational interventions. *Systems Research and Behavioral Science*, *23*(4), 451–466. https://doi.org/10.1002/sres.772

Rupp, D. E., & Beal, D. (2007). Checking in with the scientist–practitioner model: How are we doing? *The Industrial-Organizational Psychologist*, *45*(1), 35–40.

Ryan, A. M., & Ford, J. K. (2010). Organizational psychology and the tipping point of professional identity. *Industrial and Organizational Psychology*, *3*(3), 241–258. https://doi.org/10.1111/j.1754-9434.2010.01233.x

Rynes, S. L., & Gerhart, B. (1990). Interviewer assessments of applicant “fit”: An exploratory investigation. *Personnel Psychology*, *43*(1), 13–35. https://doi.org/10.1111/j.1744-6570.1990.tb02004.x

Rynes, S. L., Giluk, T. L., & Brown, K. G. (2007). The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management. *Academy of Management Journal*, *50*(5), 987–1008. https://doi.org/10.5465/AMJ.2007.27151939

Saigado, J. F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F., & Rollan, J. P. (2003). A meta-analytic study of general mental ability validity for different occupations in the european community. *Journal of Applied Psychology*, *88*(6), 1068–1081. https://doi.org/10.1037/0021-9010.88.6.1068

Sanchez, J. I., & Levine, E. L. (2009). What is (or should be) the difference between competency modeling and traditional job analysis? *Human Resource Management Review*, *19*(2), 53–63. https://doi.org/10.1016/j.hrmr.2008.10.002

Sandberg, J., & Tsoukas, H. (2011). Grasping the logic of practice: Theorizing through practical rationality. *Academy of Management Review*, *36*(2), 338–360. https://doi.org/10.5465/amr.2011.59330942

Scherbaum, C. A., Goldstein, H. W., Yusko, K. P., Ryan, R., & Hanges, P. J. (2012). Intelligence 2.0: Reestablishing a research program on g in I–O psychology. *Industrial and Organizational Psychology*, *5*(2), 128–148. https://doi.org/10.1111/j.1754-9434.2012.01419.x

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*(4), 1199–1214. https://doi.org/10.1016/j.compedu.2012.05.020

Schiepek, G., & Strunk, G. (2006). *Systemische Psychologie [Systemic psychology]*. München: Spektrum.

Schiersmann, C., & Thiel, H.-U. (2011). *Organisationsentwicklung: Prinzipien und Strategien von Veränderungsprozessen [Organizationald development: Principles and strategies of change processes]* (3rd ed.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology*, *73*(1), 46–57. https://doi.org/10.1037/0021-9010.73.1.46

Schneider, B., & Konz, A. M. (1989). Strategic job analysis. *Human Resource Management*, *28*(1), 51–63. https://doi.org/10.1002/hrm.3930280104

Schön, D. A. (1993). Generative metaphor: A perspective on problem-setting in social policy. In A. Ortony (Ed.), *Metaphor and Thought* (2nd ed., pp. 137–163). Cambridge: Cambridge University Press.

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging*, *14*(3), 483–506. https://doi.org/10.1037/0882-7974.14.3.483

Schoppek, W. (2002). Examples, rules, and strategies in the control of dynamic systems. *Cognitive Science Quarterly*, *2*, 63–92.

Schoppek, W., & Fischer, A. (2015). Complex problem solving – Single ability or complex phenomenon? *Cognitive Science*, *6*, 1669. https://doi.org/10.3389/fpsyg.2015.01669

Shepherd, A. (2001). *Hierarchial task analysis*. New York: Taylor & Francis.

Shippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B., … Sanchez, J. I. (2000). The practice of competency modeling. *Personnel Psychology*, *53*(3), 703–740. https://doi.org/10.1111/j.1744-6570.2000.tb00220.x

Shore, J., Bernstein, E., & Lazer, D. (2015). Facts and figuring: An experimental investigation of network structure and performance in information and solution spaces. *Organization Science*, *26*(5), 1432–1446. https://doi.org/10.1287/orsc.2015.0980

Shotter, J., & Tsoukas, H. (2014). In search of phronesis: Leadership and the art of judgment. *Academy of Management Learning & Education*, *13*(2), 224–243. https://doi.org/10.5465/amle.2013.0201

Siddique, C. M. (2004). Job analysis: A strategic human resource management practice. *International Journal of Human Resource Management*, *15*(1), 219–244. https://doi.org/10.1080/0958519032000157438

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, *7*(2), 268–288. https://doi.org/10.1016/0010-0285(75)90012-2

Simon, H. A. (1979). Rational decision making in business organizations. *The American Economic Review*, *69*(4), 493–513.

Simon, H. A. (1997). *Administrative behavior* (4th ed.). New York: Free Press.

Simon, H. A., & Hayes, J. R. (1976). The understanding process: Problem isomorphs. *Cognitive Psychology*, *8*(2), 165–190. https://doi.org/10.1016/0010-0285(76)90022-0

Sliter, K. A. (2015). Assessing 21st century skills: Competency modeling to the rescue. *Industrial and Organizational Psychology*, *8*(02), 284–289. https://doi.org/10.1017/iop.2015.35

Smith, G. F. (1997). Managerial problem solving: A problem-centered approach. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 371–380). Hillsdale: Erlbaum.

Snyderman, M., & Rothman, S. (1988). *The IQ controversy, the media and public policy*. Piscataway: Transaction.

Sonnenberg, H.-G. (1993). Computergestützte psychologische Diagnoseverfahren bei der Auswahl von Führungskräften [Computer-assisted psychological assessment in management selection]. *Zeitschrift Für Arbeits- Und Organisationspsychologie*, *37*(3), 146–149.

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.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24*(2), 235–270. https://doi.org/10.1086/jole.2006.24.issue-2

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, *53*, 92–101. https://doi.org/10.1016/j.intell.2015.09.005

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

Sterman, J. D. (2001). System dynamics modeling: Tools for learning in a complex world. *California Management Review*, *43*(4), 8–25. https://doi.org/10.2307/41166098

Sterman, J. D. (2006). Learning from evidence in a complex world. *American Journal of Public Health*, *96*(3), 505–514. https://doi.org/10.2105/AJPH.2005.066043

Strohschneider, S. (1986). Zur Stabilität und Validität von Handeln in komplexen Realitätsbereichen. [On the stability and validity of complex problem-solving behavior.]. *Sprache & Kognition*, *5*(1), 42–48.

Strohschneider, S. (Ed.). (2007). *Entscheiden in kritischen Situationen [Decision making in critical situations]* (2nd ed.). Frankfurt: Verlag für Polizeiwissenschaft.

Strohschneider, S., & Güss, D. (1999). The fate of the moros: A cross-cultural exploration of strategies in complex and dynamic decision making. *International Journal of Psychology*, *34*(4), 235–252.

Su, R., Golubovich, J., & Robbins, S. B. (2015). Bridging science and practice: Toward a standard, evidence-based framework of 21st century skills. *Industrial and Organizational Psychology*, *8*(02), 289–294. https://doi.org/10.1017/iop.2015.36

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 [Intelligence and complex problem solving]. *Psychologische Rundschau*, *50*(4), 220–228. https://doi.org/10.1026//0033-3042.50.4.220

Szulanski, G. (2000). The process of knowledge transfer: A diachronic analysis of stickiness. *Organizational Behavior and Human Decision Processes*, *82*(1), 9–27. https://doi.org/10.1006/obhd.2000.2884

Szulanski, G., Cappetta, R., & Jensen, R. J. (2004). When and how trustworthiness matters: Knowledge transfer and the moderating effect of causal ambiguity. *Organization Science*, *15*(5), 600–613. https://doi.org/10.1287/orsc.1040.0096

Taylor, F. W. (1911). *The principles of scientific management.* New York: Harper & Brothers.

ter Hark, M. (2010). The psychology of thinking before the cognitive revolution: Otto Selz on problems, schemas, and creativity. *History of Psychology*, *13*(1), 2–24. https://doi.org/10.1037/a0017442

Tett, R. P., & Burnett, D. D. (2003). A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, *88*(3), 500–517. https://doi.org/10.1037/0021-9010.88.3.500

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, *26*(2), 265–283. https://doi.org/10.1007/s10648-013-9243-1

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, *51*(1), 1–10. https://doi.org/10.2307/1129583

Tsoukas, H., & Dooley, K. J. (2011). Towards the ecological style: Embracing complexity in organizational research. *Organization Studies*, *32*(6), 729–735. https://doi.org/10.1177/0170840611410805

Tsoukas, H., & Hatch, M. J. (2001). Complex thinking, complex practice: The case for a narrative approach to organizational complexity. *Human Relations*, *54*(8), 979–1013.

van der Schaaf, T. W. (1993). Developing and using cognitive task typologies. *Ergonomics*, *36*(11), 1439–1444. https://doi.org/10.1080/00140139308968011

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527–579). Cambridge: MIT Press.

Vargas Cortes, M. G., & Beruvides, M. G. (1996). An analysis of middle management work in non-steady conditions. *Computers & Industrial Engineering*, *31*(1–2), 53–57. https://doi.org/10.1016/0360-8352(96)00077-0

Varghese, L., Lindeman, M. I. H., & Santuzzi, A. M. (2015). Do we really need new constructs? An argument for adapting individual predictors to dynamic environments. *Industrial and Organizational Psychology*, *8*(2), 294–301. https://doi.org/10.1017/iop.2015.37

Vester, F. (2007). *The art of interconnected thinking: Ideas and tools for a new approach to tackling complexity*. München: Malik Management.

Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Mahwah: Lawrence Erlbaum.

Vollmeyer, R., Burns, B. D., & Holyoak, K. J. (1996). The impact of goal specificity on strategy use and the acquisition of problem structure. *Cognitive Science*, *20*(1), 75–100. https://doi.org/10.1016/S0364-0213(99)80003-2

Volmerg, B., Leithäuser, T., Neuberger, O., Ortmann, G., & Sievers, B. (1995). *Nach allen Regeln der Kunst: Macht und Geschlecht in Organisationen [Power and gender in organizations]*. Freiburg: Kore.

von der Weth, R. (2001). Ich weiss, dass ich nichts weiss - Ressourcenorientiertes Handeln in komplexen Entscheidungssituationen [Resource-oriented acting in complex decision situations]. In D. Lehmann Pollheimer & H. Mey (Eds.), *Absturz im freien Fall - Anlauf zu neuen Höhenflügen: Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft* (pp. 95–112). Zürich: vdf Hochschulverlag.

Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios: Taxonomie, Entwicklung, Evaluation [Psychological diagnosis with a complex scenario: Taxonomy, development, evaluation]*. Lengerich: Pabst.

Wagner, R. K. (1991). Managerial problem solving. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 159–183). Hillsdale: Erlbaum.

Wallach, D. P. (1998). *Komplexe Regelungsprozesse: Eine kognitionswissenschaftliche Analyse [Complex control processes. A cognitive science analysis]*. Wiesbaden: DUV.

Wallach, D. P., & Tack, W. H. (1998). Wissenserwerb and Performanz bei der Regelung komplexer Systeme [Knowledge acquisition and performance in the control of complex systems]. *Kognitionswissenschaft*, *7*(3), 118–123. https://doi.org/10.1007/s001970050063

Weiß, R. H. (2006). *Grundintelligenztest Skala CFT 20-R [Culture fair intelligence test scale]* (2nd revision). Göttingen: Hogrefe.

Willke, H. (1992). Beobachtung, Beratung und Steuerung von Organisationen in systemtheoretischer Sicht [Observation, consulting, and control of organizations from a systems theory perspective]. In R. Wimmer (Ed.), *Organisationsberatung* (pp. 17–42). Wiesbaden: Gabler. https://doi.org/10.1007/978-3-322-87124-4_1

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Cognitive Science*, *4*, 58. https://doi.org/10.3389/fpsyg.2013.00058

Wirth, J., & Funke, J. (2005). Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme [Dynamic problem-solving: Construction and evaluation of a new assessment method measuring the management of complex systems]. In E. Klieme (Ed.), *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie* (pp. 55–72). Wiesbaden: VS Verlag für Sozialwissenschaften.

Wissenschaftsrat. (2015). *Zum wissenschaftspolitischen Diskurs über Große gesellschaftliche Herausforderungen [On the scientific discourse on great* *challanges for society]* (Positionspapier No. Drucksache 4594-15). Stuttgart: Wissenschaftsrat.

Wittmann, W. W., & Hattrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science*, *21*(4), 393–409. https://doi.org/10.1002/sres.653

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, *37*(1), 60–82. https://doi.org/10.1016/0749-5978(86)90044-0

Wood, R. E., Beckmann, J. F., & Birney, D. P. (2009). Simulations, learning and real world capabilities. *Education + Training*, *51*(5/6), 491–510. https://doi.org/10.1108/00400910910987273

World Economic Forum. (2016). *The future of jobs* (Global Challenge Insight Report). Cologny/Geneva: World Economic Forum.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, *40*, 1–14. https://doi.org/10.1016/j.intell.2011.11.003

Wüstenberg, S., 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

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, *30*(1), 1–53. https://doi.org/10.1086/662066

Yukl, G. A. (2010). *Leadership in organizations* (7th ed.). Upper Saddle River: Pearson.

Zaccaro, S. J., Mumford, M. D., Connelly, M. S., Marks, M. A., & Gilbert, J. A. (2000). Assessment of leader problem-solving capabilities. *The Leadership Quarterly*, *11*(1), 37–64. https://doi.org/10.1016/S1048-9843(99)00042-9

Zaccaro, S. J., Mumford, M. D., Marks, M. A., Connelly, M. S., Threlfall, K. V., Gilbert, J. A., & Fleishman, E. A. (1997). *Cognitive and temperament determinants of Army leadership* (U.S. Army Research Institute for the Behavioral and Social Sciences No. Tech. Rep. MRI 95-4). Bethesda: Management Research Institute.

Zhang, J., & Norman, D. A. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122. https://doi.org/10.1207/s15516709cog1801_3

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133. https://doi.org/10.1007/s11336-003-0974-7

Zsambok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making*. Hillsdale: Erlbaum.

List of Figures

Declaration in accordance with § 8 (1) b and § 8 (1) c of the regulations for Doktorat degrees of the Faculty of Behavioural and Cultural Studies of Heidelberg University

Promotionsausschuss der Fakultät für Verhaltens- und Empirische Kulturwissenschaften
der Ruprecht-Karls-Universität Heidelberg
Doctoral Committee of the Faculty of Behavioural and Cultural Studies, of Heidelberg University

Erklärung gemäß § 8 Abs. 1 Buchst. b) der Promotionsordnung der Universität Heidelberg
für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften
Declaration in accordance to § 8 (1) b) and § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe.
I declare that I have made the submitted dissertation independently, using only the specified tools and have correctly marked all quotations.

Erklärung gemäß § 8 Abs. 1 Buchst. c) der Promotionsordnung
der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.
I declare that I did not use the submitted dissertation in this or any other form as an examination paper until now and that I did not submit it in another faculty.

Vorname Nachname
First name Family name Jonas Neubert

Datum, Unterschrift
Date, Signature ______________________________

Appendix

Manuscript I
Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2015). Extending the assessment of complex problem solving to finite state automata: Embracing heterogeneity. *European Journal of Psychological Assessment*, *31*(3), 181–194. https://doi.org/10.1027/1015-5759/a000224

Manuscript II
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*(1), 55–69. https://doi.org/10.1016/j.intell.2015.11.004

Manuscript III
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*(2), 238–268. https://doi.org/10.1017/iop.2015.14

Original Article

Extending the Assessment of Complex Problem Solving to Finite State Automata

Embracing Heterogeneity

Jonas C. Neubert, André Kretzschmar, Sascha Wüstenberg, and Samuel Greiff

University of Luxembourg, Luxembourg

Abstract. Recent advancements in the assessment of Complex Problem Solving (CPS) build on the use of homogeneous tasks that enable the reliable estimation of CPS skills. The range of problems featured in established instruments such as MicroDYN is consequently limited to a specific subset of homogeneous complex problems. This restriction is problematic when looking at domain-specific examples of complex problems, which feature characteristics absent from current assessment instruments (e.g., threshold states). We propose to utilize the formal framework of Finite State Automata (FSA) to extend the range of problems included in CPS assessment. An approach based on FSA, called MicroFIN, is presented, translated into specific tasks, and empirically investigated. We conducted an empirical study (N = 576), (1) inspecting the psychometric features of MicroFIN, (2) relating it to MicroDYN, and (3) investigating the relations to a measure of reasoning (i.e., CogAT). MicroFIN (1) exhibited adequate measurement characteristics and multitrait-multimethod models indicated (2) the convergence of latent dimensions measured with MicroDYN. Relations to reasoning (3) were moderate and comparable to the ones previously found for MicroDYN. Empirical results and corresponding explanations are discussed. More importantly, MicroFIN highlights the feasibility of expanding CPS assessment to a larger spectrum of complex problems.

Keywords: complex problem solving, finite state automata, MicroFIN, MicroDYN, multitrait-multimethod

Due to their relevancy for dealing with the challenges of our times, an individual's skills in coping with complex problems are generally considered as part of the so called 21st century skills (cf. Griffin, McGaw, & Care, 2012). Scientifically, these skills in coping with complex problems are investigated under the label of Complex Problem Solving (CPS; e.g., Buchner, 1995; Sternberg & Frensch, 1991), sometimes also called Dynamic Problem Solving (e.g., Greiff, Wüstenberg, & Funke, 2012), Dynamic Decision Making (e.g., Brehmer, 1992), or Interactive Problem Solving (e.g., OECD, 2013; see Greiff et al., 2013, for a discussion of different names). In the following, we will use the term CPS to refer to the construct.

Recently, CPS has been added to a range of high-profile studies as representatives of domain-general and transversal skills, for example to the Programme for International Student Assessment in its 2012 cycle (PISA; OECD, 2013), the arguably most important large-scale assessment world wide.

One of the defining characteristics of the complex problems employed within CPS assessment in large-scale efforts and in general are their changes in reaction to user interaction and/or passage of time. Additionally, they demand active interventions and feature a multitude of interrelated factors (see Buchner, 1995). Dealing successfully with such complex problems has been shown to be related to, but separable from other (cognitive) constructs, such as reasoning ability and working memory capacity, both, conceptually and empirically (e.g., Schweizer, Wüstenberg, & Greiff, 2013; Sonnleitner, Keller, Martin, & Brunner, 2013; Wüstenberg, Greiff, & Funke, 2012; see also Bühner, Kröner, & Ziegler, 2008; Wittmann & Süß, 1999).

On the side of the problem solver, the dynamic, interactive, and complex problems result in characteristic requirements with regard to processes of knowledge acquisition and knowledge application (for details see Fischer, Greiff, & Funke, 2012; Osman, 2010). An individual's skills to deal with such problems are called CPS skills and within this study we aim at broadening the view on their assessment.

Assessment of Complex Problem Solving

Looking for a viable assessment of CPS, essential elements are the utilization of formal frameworks and the application of several multiple complex systems within these formal

DOI: 10.1027/1015-5759/a000224

frameworks (cf. Greiff et al., 2012). Formal frameworks allow for the analysis of the underlying structure of problems instead of surface features and hence, the systematic examination and comparison of assessment instruments (Funke, 2001; Greiff et al., 2012). Multiple complex systems built on a formal framework and employed within one assessment instrument are a viable and straightforward way to meet the psychometric requirement of stochastically independent indicators to estimate the target constructs (cf. Greiff et al., 2012; Wüstenberg et al., 2012).

Currently, CPS assessment instruments combining formal frameworks and multiple complex systems are mainly based on one specific formal framework: Linear Structural Equations (LSE).[1] In LSE, problems are formally described as a set of linear equations of quantitative variables. That is, the problems are formalized as several input and output variables that are connected by linear relations between each other (e.g., the amount of different fertilizers changing the growth of flowers). Within the assessment, participants have to explore several of these relations by manipulating the input variables, observing the resulting changes in the output variables and, from this, deriving the causal relations between input and output variables (knowledge acquisition; cf. Novick & Bassok, 2005).

Subsequently, participants have to use their knowledge to reach target values in the output variables (knowledge application).

Historical instances of instruments targeting CPS with the help of LSE featured a singular complex problem (e.g., MultiFlux; Bühner et al., 2008) and consequently suffered from psychometric problems due to one-item-testing (Greiff et al., 2012). To overcome this problem, CPS assessment nowadays employs multiple problems in succession and is consequently able to reliably estimate a persons CPS skills. Examples of instruments combining LSE as formal framework with multiple complex systems are *Genetics Lab* (Sonnleitner et al., 2012), *MicroDYN* (e.g., Greiff et al., 2012), and the assessment of CPS in PISA 2012, that is partially based on MicroDYN (OECD, 2013).

The utilization of the framework of LSE in combination with multiple complex systems has resulted in reliable instruments targeting core features of CPS with a focus on psychometrics requirements (e.g., Greiff & Wüstenberg, 2014). But when comparing the range of problems employed in assessment with the original breadth of the construct, instruments relying on LSE and, thus, on quantitative relations between a set of variables, are necessarily restricted as they have to follow a predefined, rather narrow pattern of relations between elements (cf. Funke, 2010). Consequently, the problems presented to participants are rather homogeneous with regard to the kinds of relations that have to be explored.

Complex problems found in specific domains, on the other hand, for example within the domain of chemistry, include features that are absent from current CPS assessment based on LSE. Examples are relations between input and output variables, which feature strong interactions between input variables as found in Le Chatelier's principle of chemical equilibria (e.g., Scherer & Tiemann, 2012) and qualitative changes in the problem after reaching a threshold (e.g., water freezing to ice). These and other features not present in current LSE-based CPS assessment are also part of historical notions of CPS (e.g., Dörner, Kreuzig, Reither, & Stäudel, 1983; Funke, 2010) and real-world examples of complex problems (e.g., Ackerman, 1992).

As a consequence, LSE-based assessment instruments are currently unable to assess differences and commonalities with regard to a range of problem characteristics and corresponding requirements on the side of the problem solver. To give an example: Vary-one-thing-at-a-time (VOTAT; Tschirgi, 1980) can be considered an sufficient and optimal strategy in the LSE-based tasks mentioned above (Vollmeyer, Burns, & Holyoak, 1996; see Rollett, 2008, for a comprehensive discussion of strategies in LSE-based tasks), but not in the case of problems featuring threshold or equilibrium states (cf. McElhaney & Linn, 2011). Hence, we have to draw conclusions on general exploration behavior in complex problem situations based on one specific strategy (i.e., VOTAT) when using LSE-based tasks, whereas broader sets of strategies and an adaptive use of them are necessary to deal with complex problems in general (see Levy & Wilensky, 2011, for an example of adaptive exploration strategies).

To counter the limitation in the breadth of problems included in instruments and facilitate a greater heterogeneity in tasks, we propose to expand CPS assessment to include tasks based on another formal framework introduced by Buchner and Funke (1993), Finite State Automata (FSA).

The Formal Framework of Finite State Automata

In FSA, problems are formalized as predefined states with transitions between these states triggered by events such as user interactions or passing of time. In contrast to LSE, the type of relation formalized as transition does not have to follow a specific pattern (i.e., quantitative relations). Consequently, FSA-based CPS assessment can easily include a variety of features that overcome the current homogeneity of LSE-based tasks, for example by including complex problems that require different exploration strategies. That is, the use of FSA allows for the inclusion of heterogeneous causal relations in CPS assessment, thereby leading to a broader assessment of CPS skills.

At the same time, FSA-based CPS assessment can preserve advantages of established instruments based on LSE such as multiple tasks comparable by their causal structure and a general layout resembling that of established instruments (e.g., relating input variables to output variables). Additionally, FSA-based tasks were already utilized

[1] Please note that researchers focusing on basic human processes and a modeling of complex problems following real-world examples are successfully utilizing a much broader range of complex problems (see for example the work by Gonzalez, Vanyukov, & Martin, 2005). For the sake of brevity, we focus on instruments directed toward skill assessment here.

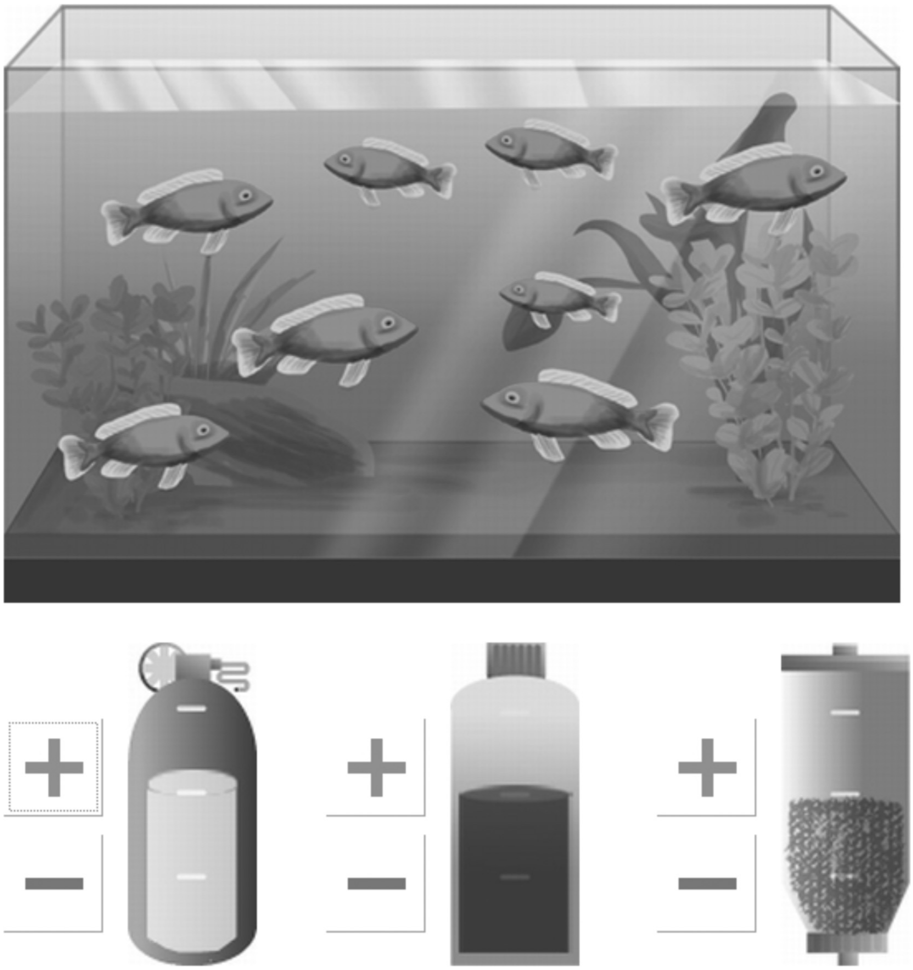

Figure 1. Screenshot of the MicroFIN task Fish-o-mat. The three containers at the bottom represent the input variables, the aquarium on top the output variable. Input variables' values can be varied by pressing the plus and minus signs next to them. Each of the input variables features four input values (empty, low, medium, high).

in laboratory studies targeting CPS, providing elaborations on the FSA framework itself and ways of mathematical and abstract representation (e.g., Buchner & Funke, 1993; Funke, 2001). Outside the realm of CPS research, FSAs are commonly used to formalize the workings of a broad range of appliances, for example for the purpose of programming vending machines, combination locks, and turnstiles (e.g., Anderson, 2006; Rich, 2008).

Empirical applications combining FSA and multiple complex systems, have not surfaced, yet (for a notable attempt, see the PISA framework for Problem Solving in 2012; OECD, 2013, 2014). To empirically investigate the utility of the combination of FSA with the approach of multiple complex systems, we therefore developed a set of tasks that represent different options exceeding the limitations of LSE-based instruments. The resulting approach is called MicroFIN, a label coined by Greiff and Funke (2009; see also Greiff, Fischer, et al., 2013).

Like other assessment vehicles and along the theoretical understanding of CPS (e.g., MultiFlux, Bühner et al., 2008; MicroDYN, Greiff et al., 2012; Genetics Lab, Sonnleitner et al., 2012), MicroFIN tasks are structured into two separate phases: Tasks begin with the (1) knowledge acquisition phase, where participants can freely explore the simulation, followed by an assessment of the acquired knowledge. In the following (2) knowledge application phase, participants are then asked to reach given target states, assessing their capabilities in applying the acquired knowledge.

An illustration of a MicroFIN task, called "Fish-o-maton," is depicted in Figure 1, showing the layout of the task as presented to participants. The corresponding state-transition diagram, an abstract formal representation of the underlying finite state automaton can be found in Figure A1 in the Appendix.

The "Fish-o-maton" includes three ordinal input variables, each with four input values and an output variable visualized as an aquarium (nominal variable with ordinal elements).

The output variable has five possible output values: Empty, soiled, few fish, moderate number of fish, and many fish. Whereas the general structure of the task is similar to LSE-based items, with input variables related to output variables, the specific relation between input and output variables is unfeasible to implement within LSE: Transitions are triggered when all input variables are moved into equivalent states and out of them. To give an example, if a participant brings all input variables to a low amount of input, few fish are displayed in the aquarium. If she moves the first input variable out of this equilibrium by increasing its value, the output variable changes to a display of a soiled aquarium, that is, a qualitative change in the output variable not feasible within LSE-based instruments.

Consequently, an exploration behavior following the application of VOTAT, which would be successful in the case of LSE-based tasks, that is, varying the inputs in isolation from zero to three, would lead to the omission of the central states of the "Fish-o-maton." Hence, the task is an example of expanding the range of necessary exploration behavior, while maintaining general features of established CPS assessment instruments.

Table 1 gives an overview of the features of all five utilized MicroFIN tasks. As explicated above for the "Fish-o-maton," the table includes information on the specific kind of relations between input and output variables realized in each task, differentiating the tasks from each other, as well as from established instruments based on LSE such as MicroDYN. Furthermore, a detailed description of all utilized MicroFIN tasks is given in the Appendix.

Focus of the Empirical Study

The inquiries of the empirical study are directed toward the exploration of the empirical features of MicroFIN. Specifically, attention is directed toward three main directions.

(1) *Securing psychometric properties and a measurement model for MicroFIN*

A solid psychometric foundation is essential for assessment. Therefore, proficient levels of reliability and adequate item difficulty are expected when assessing the dimensions of CPS (Hypothesis 1a). With regard to these dimensions of CPS, a well-proven differentiation in CPS research is anticipated (e.g., Funke, 2001) with separable empirical dimensions for (1) knowledge acquisition and (2) knowledge application (Hypothesis 1b).

(2) *Relating MicroFIN to an established instrument of CPS assessment*

The latent dimensions of CPS as assessed via MicroFIN are expected to capture the same underlying constructs as established instruments based on LSE, pointing to convergent validity (i.e., knowledge acquisition and knowledge application). Therefore, a latent multitrait-multimethod model is expected to hold with instruments based on FSA and LSE loading on the same latent variables. Again,

Table 1. Features of the utilized MicroFIN tasks

Task	No. of input variables[a]	No. of states[b]	Input functioning[c]	Output functioning[c]	Other features[d]
Task 1: Concert-o-maton	2 × 2 × 3	12	Quantitative (ordinal) and qualitative (dichotomous)	Quantitative (ordinal)	Opposite subsystems
Task 2: Plan-o-maton	4 × 4 × 4 × 4	24	Qualitative (nominal)	Qualitative (dichotomous)	Relative positioning
Task 3: Plant-o-maton	4 × 4 × 4	64	Quantitative (ordinal)	Qualitative (nominal)	Threshold value
Task 4: Fish-o-maton	4 × 4 × 4	64	Quantitative (ordinal)	Quantitative (ordinal) and Qualitative (nominal)	Equivalence
Task 5: Flooz-o-maton	3 × 3 × 3	27	Quantitative (ordinal) and qualitative (nominal)	Quantitative (ordinal) and Qualitative (nominal)	Start button and threshold values

Notes. Table intended to give an overview of the features varied between MicroFIN tasks.
[a]Number of input variables and possible values per variable. Example: 2 × 3 means one variable with two possible values and another variable with three possible values.
[b]Number of different states, the finite state automaton can be in (i.e., combinations of input and output variables).
[c]Input/Output variables are about qualitative (i.e., dichotomous or nominal) or quantitative (i.e., ordinal) variation, or combinations of both.
[d]Other features of the task, representing unique characteristics; Opposite subsystems: The task features two subsystems with opposite reaction to input variation. Relative positioning: Relative position of subelements as relevant input for transitions between output states. Start button: Effects of input manipulation are only observable after pressing a button (comparable to MicroDYN). Threshold values: The tasks behavior changes if one variable is put above a certain level. Equivalence: The task features equilibrium states as relevant input. For further information please refer to the detailed description of tasks in the Appendix. The commonalities between task features as shown in the table are considerably lower in FSA than for tasks based on LSE, where a comparable line in Table 1 for *all* items would read: Number of input variables ranging from two to four with up to 20 values per input variable (i.e., finer granulation in inputs). Quantitative inputs and outputs and a start button as other feature.

separate dimensions for knowledge acquisition and knowledge application as indicated by both instruments are expected. Additionally, method factors for both dimensions in MicroFIN are anticipated, emphasizing the specific requirements resulting from the broader range of problems implemented in MicroFIN (Hypothesis 2).

(3) *Exploring the relations to reasoning*

An important concept to check the instrument against in terms of discriminant validity is reasoning ability (cf. Sonnleitner et al., 2013; Wüstenberg et al., 2012). Extending the findings of Wüstenberg et al. (2012) for an LSE-based instrument, a positive relation but also separability is expected between reasoning and the dimensions of CPS as measured by the multitrait-multimethod model introduced in Hypothesis 2 (Hypothesis 3).

Materials and Methods

To empirically test these hypotheses, we conducted a study in the educational context, relating MicroFIN tasks to MicroDYN, an established instrument targeting CPS and based on LSE (Greiff et al., 2012). To inquire the relations to reasoning we employed the nonverbal scale of the Cognitive Abilities Test (CogAT; German adaptation by Heller & Perleth, 2000). All instruments were fully computer-based.

Participants

576 German high school students (262 males) attending grades 8–12 participated in the study (age between 13 and 18, $M = 14.95$, $SD = 1.30$). Participants attended one of three school tracks within the same school, together covering all educational tracks in German high schools. Participation was voluntary and we received informed consent from parents. With regard to incentives, participants received monetary compensation on a per-class basis that was given to the class inventory. Data of two participants had to be excluded from analyses due to technical problems during assessment.

Measures

Complex Problem Solving

MicroFIN

The features of the set of five MicroFIN tasks employed in this study, as well as a detailed description of the task "Fish-o-maton" has already been given (see also Table 1 and the Appendix).[2] Instructions on MicroFIN included an additional trial task that was excluded from analysis. Generally, all MicroFIN tasks feature separate phases for the exploration of the problem's relations and the assessment of (1) knowledge acquisition and (2) knowledge application, each dimension targeted with the help of specific items. Overall, each MicroFIN task takes approximately 5 min to complete.

More specifically, after participants freely explore the task, items in the knowledge acquisition phase target the gathered knowledge. Within each item, an outcome state and a transition is given, and the initial state of the task needs to be selected by the participant (constructed response items, cf. Buchner, 1995). In the example of the "Fish-o-maton," such items would present a specific manipulation of an input variable (e.g., raising the level of the first input variable from two to three) and the resulting state of the task (e.g., a soiled aquarium, with the input of the remaining two input variables being two). Participants are then asked to construct the initial state of the "Fish-o-maton" prior to the manipulation from given elements (e.g., all inputs being on the level of two and an aquarium displaying a medium amount of fish). The item type was successfully introduced to CPS research in laboratory applications of FSA-based tasks, making them a natural choice to assess participants' knowledge (e.g., Buchner & Funke, 1993). Two constructed response items are included per task.

In the knowledge application phase, items are featuring the task in a specific state and participants are asked to manipulate the input variables (i.e., trigger transitions) to reach a goal state that is presented visually and verbally at the beginning of each item (e.g., a specific amount of fish). Two of these items are included per task.

Scoring of knowledge acquisition: Participants receive credit for correctly constructing the initial state of an item and no credit if they fail to do so. A sum score over the two items ranging from 0 to 2 per task is utilized as a manifest indicator in latent analyses to reduce the number of estimated parameters.

Scoring of knowledge application: Participants receive credit for reaching the target state. No credit is given, when participants fail to reach the target state. Again, a sum score for both knowledge application items is used as manifest indicator with a range of 0 to 2 for each MicroFIN task.

MicroDYN

The MicroDYN approach assesses CPS based on the formal framework of LSE and has been covered extensively elsewhere (e.g., Greiff & Funke, 2009; Greiff et al., 2012; Greiff & Wüstenberg, 2014; Schweizer et al., 2013; Wüstenberg et al., 2012). Tasks are defined by up to three quantitative input variables relating to one or

[2] The MicroFIN tasks utilized in the study can be obtained from the first author upon request for academic purposes.

several output variables in a linear quantitative way. The connections between these variables have to be discovered and, subsequently, used to reach target values. The distinction between (1) knowledge acquisition and (2) knowledge application implemented in MicroFIN can also be found in MicroDYN.

Scoring of both phases of CPS assessment follows the recommendation of Wüstenberg et al. (2012), with credit given for correct models and no credit for wrong models in knowledge acquisition. In knowledge application, credit is given for reaching the target values and no credit otherwise. There were eight MicroDYN tasks employed in the study, each taking an average of 5 min to complete, plus an instructional task that was excluded from analysis. The underlying linear equations of the eight MicroDYN tasks can be found in Table A2 in the Appendix.

Reasoning

For the reasoning assessment, participants completed a computer-adapted version of the nonverbal scale of the Cognitive-Abilities Test (CogAT; Heller & Perleth, 2000). There, participants are asked to identify a figure, completing a 3×3 matrix of figures related by combination rules (similar to Ravens Advanced Progressive Matrices, Raven, Raven, & Court, 1998). Credit is given for correct solutions and no credit for wrong answers. Two items (item 9 and 14) of the 25 items included in the CogAT have been shown to be insolvable due to ambiguous solutions (Segerer, Marx, & Marx, 2012) and were excluded from analysis.

Statistical Analysis

Statistical analyses were conducted in MPlus Version 7 (Muthén & Muthén, 2012) utilizing descriptive analyses (Hypothesis 1a), confirmatory factor analysis (Hypothesis 1b), and structural equation modeling for the estimation of latent factors and their relations (Hypotheses 2 and 3). The empirical evaluation of constructs and instruments was done within multitrait-multimethod models (cf. Eid, Lischetzke, Nussbeck, & Trierweiler, 2003) allowing for the estimation of method specific effects for instruments targeting the same construct. As our indicators are ordered categorical variables, we used weighted least squares with means and variances adjusted estimation (WLSMV; Muthén, du Toit, & Spisic, 1997). Global goodness-of-fit was evaluated by χ^2-tests, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), and Tucker Lewis Index (TLI).

According to Hu and Bentler (1999), a χ^2 to *df* ratio < 2 and RMSEA values $\leq .06$ indicate a good global fit, as do values $\geq .95$ for CFI and TLI. For the comparison of models, we utilized a specific procedure for WLSMV estimation integrated in MPlus to compute χ^2-difference values (Muthén & Muthén, 2012, p. 451).

Results

Hypothesis 1a: Psychometric Properties of MicroFIN

Descriptive analyses for MicroFIN were the basis for analyzing item difficulties and reliability for both phases of CPS as targeted in Hypothesis 1a. In knowledge acquisition, the average success rate of participants (i.e., item difficulty) ranged from $p = .18$ to $.49$ ($M = .34$, $SD = 0.14$) across the five tasks. In knowledge application item difficulty ranged from $p = .63$ to $.77$ ($M = .70$, $SD = 0.06$), that is, items targeting knowledge application were easier compared to items targeting knowledge acquisition. Reliability for both dimensions was acceptable with McDonalds's omega (Zinbarg, Revelle, Yovel, & Li, 2005) of $\omega = .79$ and $\omega = .78$ for knowledge acquisition and knowledge application, especially when taking the number of only five MicroFIN tasks into account. The results supported Hypothesis 1a, confirming adequate item difficulty and reliability for MicroFIN.

Hypothesis 1b: A Measurement Model for MicroFIN

As laid out under Hypothesis 1b, we expected a 2-dimensional measurement model differentiating between knowledge acquisition and knowledge application for MicroFIN. Results indicated a good model fit for this model ($\chi^2(34) = 52.684$, $p = .021$, RMSEA = 0.031, CFI = 0.989, TLI = 0.986) with a latent correlation between dimensions pointing to strongly related, but nonetheless separable dimensions of CPS as measured by MicroFIN ($r = .81$, $p < .001$).

An alternative measurement model conflating the two dimensions resulted in significantly worse model fit ($\chi^2(35) = 90.118$, $p < .001$, RMSEA = 0.053, CFI = 0.968, TLI = 0.958, $\chi^2\Delta(1) = 24.398$, $p < .001$). Therefore, Hypothesis 1b, which assumed a measurement model with separate dimensions for knowledge acquisition and knowledge application for MicroFIN, was supported.

Descriptives and Measurement Models for the Remaining Instruments

MicroDYN

For the eight MicroDYN tasks item difficulty for knowledge acquisition was in the range of $p = .08$ and $.68$ ($M = .42$, $SD = 0.26$) and $p = .05$ to $.50$ ($M = .33$, $SD = 0.14$) for knowledge application. Item difficulties were comparable to the ones reported in other applications of MicroDYN (e.g., Schweizer et al., 2013). Reliability was excellent with $\omega = .93$ (knowledge acquisition) and $\omega = .88$ (knowledge application).

Replicating previous findings (e.g., Schweizer et al., 2013; Wüstenberg et al., 2012), a 2-dimensional model with separate dimensions for knowledge acquisition and knowledge application resulted in the best fitting measurement model for MicroDYN ($\chi^2(103) = 225.982$, $p < .001$, RMSEA = 0.047, CFI = 0.979, TLI = 0.976; latent correlation of dimensions $r = .82$, $p < .001$).

CogAT

For the CogAT, item difficulty ranged from $p = .41$ to .81 ($M = .67$, $SD = 0.12$) and reliability can be considered good with $\omega = .95$. Due to the large number of 23 items, we used item-to-construct parceling according to Little, Cunningham, Shahar, and Widaman (2002) to reduce the number of estimated parameters. Three parcels were constructed with comparable average loading of the items on the parcels between $M\lambda = .64$ and .66 and mean item difficulty ranging from $Mp = 0.67$ to .70. An essentially tau-equivalent measurement model (e.g., Novick, 1966) assuming equal loadings for parcels, but allowing for differences in intercepts fitted the data well ($\chi^2(2) = 3.908$, $p = .142$, RMSEA = 0.042, CFI = 0.998, TLI = 0.997).

Hypotheses 2 and 3: Structural Models

Complex Problem Solving (CPS)

Hypothesis 2 was directed toward the relation of MicroFIN and MicroDYN and the question, whether both instruments empirically target the same construct. Figure 2 shows the multitrait-multimethod model utilized to test the hypothesis, a correlated trait-correlated method minus one model (CT-C(M-1); Eid et al., 2003), including the same latent dimensions measured by both instruments and method factors for the number of methods minus one (i.e., one method served as reference method). This way, both instruments were assumed to target the same dimensions of CPS, namely knowledge acquisition and knowledge application, while specific aspects of MicroFIN resulting from the broader range of included problem features being explicitly modeled via method factors for both dimensions of CPS. MicroDYN as the established instrument served as the reference method. The resulting model fitted the data well ($\chi^2(285) = 445.247$, $p < .001$, RMSEA = 0.031, CFI = 0.980, TLI = 0.978). Results of alternative structural models with latent factors for both dimensions of CPS and both instruments estimated separately can be found at: http://dx.doi.org/10.6084/m9.figshare.1160505.

Similar to the results for the separate measurement models of MicroFIN and MicroDYN, the latent indicators for knowledge acquisition and knowledge application correlated significantly in the CT-C(M-1) model ($r = .82$, $p < .001$, see Figure 2). Both method factors for MicroFIN were significantly correlated, pointing toward a somewhat generalized method effect ($r = .54$, $p < .001$, see Figure 2). Loadings of individual MicroFIN and MicroDYN items on the latent factors of CPS and on the method factors for the MicroFIN items can be found in Table A1 in the Appendix. An alternative model (not presented) with MicroFIN as the reference method yielded comparable results with regard to correlations and loadings. In summary, Hypothesis 2, which assumed the assessment of the same constructs by both instruments, was supported, with method specific factors differentiating their assessment through MicroFIN from MicroDYN.

Reasoning

Targeting Hypothesis 3, we included a separate latent factor for reasoning in the multitrait-multimethod model described above (see also Figure 2). The resulting model fitted the data well ($\chi^2(361) = 540.359$, $p < .001$, RMSEA = 0.029, CFI = 0.978, TLI = 0.976) and featured significant correlations between knowledge acquisition and reasoning ($r = .63$, $p < .001$) and knowledge application and reasoning ($r = .60$, $p < .001$). The method factors of MicroFIN were also significantly correlated with reasoning (knowledge acquisition, $r = .34$, $p < .001$, knowledge application, $r = .28$, $p < .001$).

The correlations of the CPS dimensions with reasoning did not indicate identity of the two concepts: For knowledge acquisition the 99% confidence interval of the correlation was [.55, .71], for knowledge application [.50, .71], both of them far and significantly different from values near one. Again, utilizing MicroFIN as the reference method led to similar results (not presented). In summary, the results supported Hypothesis 3, pointing to a moderate relation between the dimensions of CPS and reasoning but not identity and an influence of reasoning on the method specific factors.

Discussion

The extension of CPS assessment to include multiple tasks formalized as finite state automata led to a broader range of problem features to be included in our assessment, thereby overcoming the homogeneity of current instruments building on LSE-tasks (see Table 1 and the detailed description of tasks in the Appendix).

(1) *Securing psychometric properties and a measurement model for MicroFIN*
The reliability of MicroFIN naturally decreased with the range of included features not shared between tasks and the rather low number of tasks. However, adequate reliability could still be achieved for both dimensions of CPS. Together with the reasonable item difficulties, these findings highlight the empirical viability of including a broader range of problem features in CPS assessment via MicroFIN (Hypothesis 1a supported). On the other hand, the results also underline the benefits in terms of reliability when utilizing homogeneous tasks in assessment as in MicroDYN

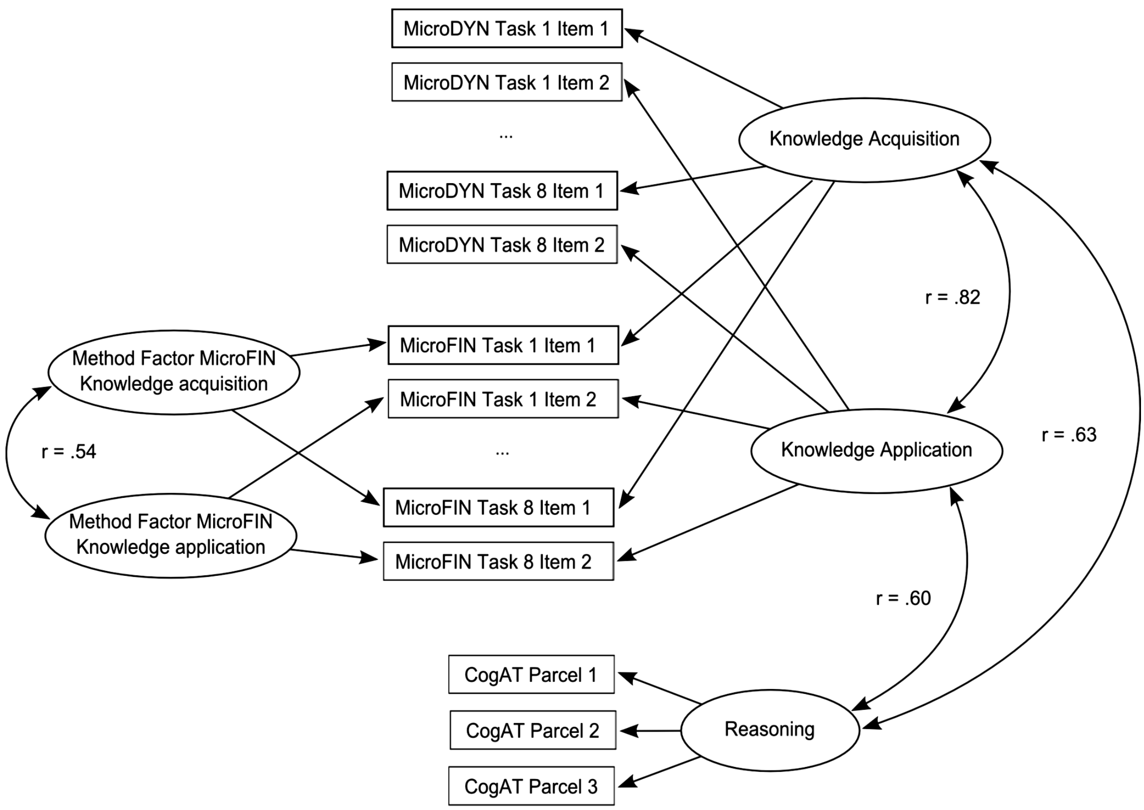

Figure 2. CT-C(M-1) multitrait-multimethod model. The model depicts the latent relations between the dimensions of CPS, knowledge acquisition and knowledge application, two method factors modeling specific aspects of MicroFIN for both dimensions, and a latent factor indicating participants' reasoning ability. Latent correlations for the method factors, error variances, and manifest indicators for some MicroFIN and MicroDYN tasks were omitted from the figure for better accessibility. The factor loadings of the model are available in Table A1 in the Appendix.

(e.g., Greiff & Wüstenberg, 2014). With regard to dimensionality, empirical results indicated the separability of knowledge acquisition and knowledge application in MicroFIN, thereby supporting our expectations of an assessment allowing for the differentiation of both theoretically derived dimensions of CPS (Hypothesis 1b supported).

(2) *Relating MicroFIN to an established instrument of CPS assessment*
Combining MicroFIN with MicroDYN showed the expected proximity of both instruments when analyzed with the help of a multitrait-multimethod model. The approaches indicated the same latent traits, with method effects differentiating both instruments and their specific features (Hypothesis 2 supported). While MicroFIN generally targeted the same skills as MicroDYN, we also saw specific problem features realized within MicroFIN leading to differently accentuated requirements (e.g., a different set of necessary exploration strategies as in the case of the ''Fish-o-maton''). This finding is represented in the substantial loadings on the method factors for MicroFIN. Looking at the heterogeneity of the underlying construct and the general goal of this study, the expansion of heterogeneity in assessment instruments to new problem relations with high real-world importance, method effects like these come as no surprise. On the contrary, they show the need for a broader assessment of CPS that also includes heterogeneous sets of tasks not covered within LSE-based instruments alone. We would expect these additional elements included in MicroFIN and represented in the method factors to be generally useful, for example by providing additional value when predicting real- life indicators of successful problem solving.

(3) *Exploring the relations to reasoning*
he correlations between the dimensions of CPS as resulting from the multitrait-multimethod model and the CogAT as an indicator for reasoning ability were significant and in the range reported for MicroDYN by Wüstenberg et al. (2012). Both method factors of MicroFIN were also significantly correlated to reasoning, even though the relation was weaker than the relation between reasoning and the dimensions of CPS. That is, the additional requirements introduced by the more heterogeneous assessment of CPS were also

associated with reasoning, as could be expected, but not exclusively due to a higher influence of reasoning in MicroFIN. In summary, both dimensions of CPS and the method factors of MicroFIN showed substantial relations to reasoning, but not identity (Hypothesis 3 supported).

Naturally, there are also limitations to take into account when considering the results of our study. Bearing in mind the enormous possibilities of the FSA framework, the sample of MicroFIN tasks was not representative of either restrictions or focus of FSA-based CPS assessment in general. We combined the framework of FSA with the approach of multiple complex systems, highlighting some first steps of how to exceed the margins of current LSE- based CPS assessments (see Table 1) and establishing the empirical feasibility of such an endeavor, but an exhaustive analysis of different problem features was beyond the scope of this study. Furthermore, an analysis of exploration strategies, necessary in response to the broader range of problem features implemented in MicroFIN, was not included here (see for example Rollett, 2008, for a comprehensive account on strategies in LSE-based tasks). The analyses of participants' process data certainly offers ample opportunity to dive into differences in terms of underlying process requirements in response to differently structured complex problems of various sizes (e.g., featuring a variety of possible states and problem features). And the availability of heterogeneous problem features combined in one approach of assessment might highlight interindividual differences in the adaptability of strategies across differently structured problems (see McElhaney & Linn, 2011). Future studies will also have to show whether the additional requirements introduced by MicroFIN allow for a better prediction of external outcomes of successfully handling complex problems outside of assessment contexts and how MicroFIN compares to the broader range of complex problems utilized in research (i.e., also in relation to instruments not focusing on assessment, e.g., Gonzalez, Vanyukov, & Martin, 2005). Finally, the question remains to be answered whether the semantic embedding used in MicroFIN represents a problem by triggering the formation of hypotheses that are not systematically tested during exploration (cf. Beckmann & Goode, 2013). Effects of semantic embedding as identified by Beckmann and Goode (2013) for LSE-based tasks and potential ways to minimize them remain to be researched for MicroFIN and other instruments building on FSA.

Looking at the assessment of CPS, MicroFIN represents the expansion of instruments to a formal framework that was previously limited to applications in laboratory settings, namely finite state automata. The framework is opening up the possibility of formalizing a broad range of problem features previously excluded from CPS assessment, thereby paving the way toward more heterogeneity of complex problems included in large-scale assessments, such as PISA or national school monitoring efforts. And by combining the formal framework of finite state automata with the use of multiple complex systems, the MicroFIN approach is facilitating a CPS assessment overcoming the current homogeneity of tasks building on LSE, while maintaining their advantage of a psychometrically sound foundation. In this paper we established the empirical applicability of the MicroFIN approach in a context of assessment and our findings give hope for a CPS assessment reaching out to a more valid reflection of the complex problems we encounter in the real world.

References

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology, 77*, 598–614.

Anderson, J. A. (2006). *Automata theory with modern applications*. Cambridge, MA: Cambridge University Press.

Beckmann, J. F., & Goode, N. (2013). 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*, 1–20.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*, 211–241.

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* (pp. 27–63). Hillsdale, NJ: Erlbaum.

Buchner, A., & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. *The Quarterly Journal of Experimental Psychology, 46*, 83–118.

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

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

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods, 8*, 38–60.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving, 4*, 19–42.

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

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

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

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

Greiff, S., & Funke, J. (2009). Measuring complex problem solving: the MicroDYN approach. In F. Scheuermann & J. Björnsson (Eds.), *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing* (pp. 157–163). Luxembourg: Office for Official Publications of the European Communities.

Greiff, S., & Wüstenberg, S. (2014). Assessment with microworlds using MicroDYN: Measurement invariance and latent mean comparisons: Psychometric properties across several student samples and blue-collar workers. *European Journal of Psychological Assessment*. Advance online publication. doi: 10.1027/1015-5759/a000194

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

Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013). Complex problem solving in educational contexts – something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology, 105*, 364–379.

Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. New York, NY: Springer.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision* [Cognitive Abilities Test (CogAT; Thorndike, L. & Hagen, E., 1954–1986) – German adapted version]. Göttingen, Germany: Beltz-Test.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55.

Levy, S. T., & Wilensky, U. (2011). Mining students' inquiry actions for understanding of complex systems. *Computers & Education, 56*, 556–573.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 151–173.

McElhaney, K. W., & Linn, M. C. (2011). Investigations of a complex, realistic task: Intentional, unsystematic, and exhaustive experimenters. *Journal of Research in Science Teaching, 48*, 745–770.

Muthén, B. O., du Toit, S. H. C., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes.* Unpublished technical report. Retrieved from http://pages.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Novick, L. R., & Bassok, M. (2005). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 321–349). New York, NY: Cambridge University Press.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology, 3*, 1–18.

OECD. (2013). *PISA 2012 assessment and analytical framework*. Paris, France: Organisation for Economic Cooperation and Development.

OECD. (2014). *PISA 2012 Results: Creative problem solving (Volume V)*. Paris, France: OECD Publishing.

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

Raven, J., Raven, J. C., & Court, J. H. (1998). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.

Rich, E. (2008). *Automata, computability and complexity: Theory and applications*. Upper Saddle River, NJ: Prentice Hall.

Rollett, W. (2008). *Strategieeinsatz, erzeugte Information und Informationsnutzung bei der Exploration und Steuerung komplexer dynamischer Systeme* [Use of strategy, generated information and use of information when exploring and controlling complex dynamic systems]. Berlin, Germany: Lit.

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.

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.

Segerer, R., Marx, A., & Marx, P. (2012). Unlösbare Items im KFT 4–12+R [Unsolvable items in the CFT 4–12+R]. *Diagnostica, 58*, 45–50.

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*, 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*, 289–305.

Sternberg, R. J., & Frensch, P. A. (1991). *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Erlbaum.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*, 1–10.

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

Wittmann, 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. 77–108). Washington, DC: American Psychological Association.

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

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*, 123–133.

Date of acceptance: April 3, 2014
Published online: October 17, 2014

Jonas C. Neubert

University of Luxembourg
Educational Measurement and Applied Cognitive Science (EMACS)
6, rue Richard Coudenhove-Kalergi
1359 Luxembourg
Luxembourg
E-mail jonas.neubert@uni.lu

Appendix

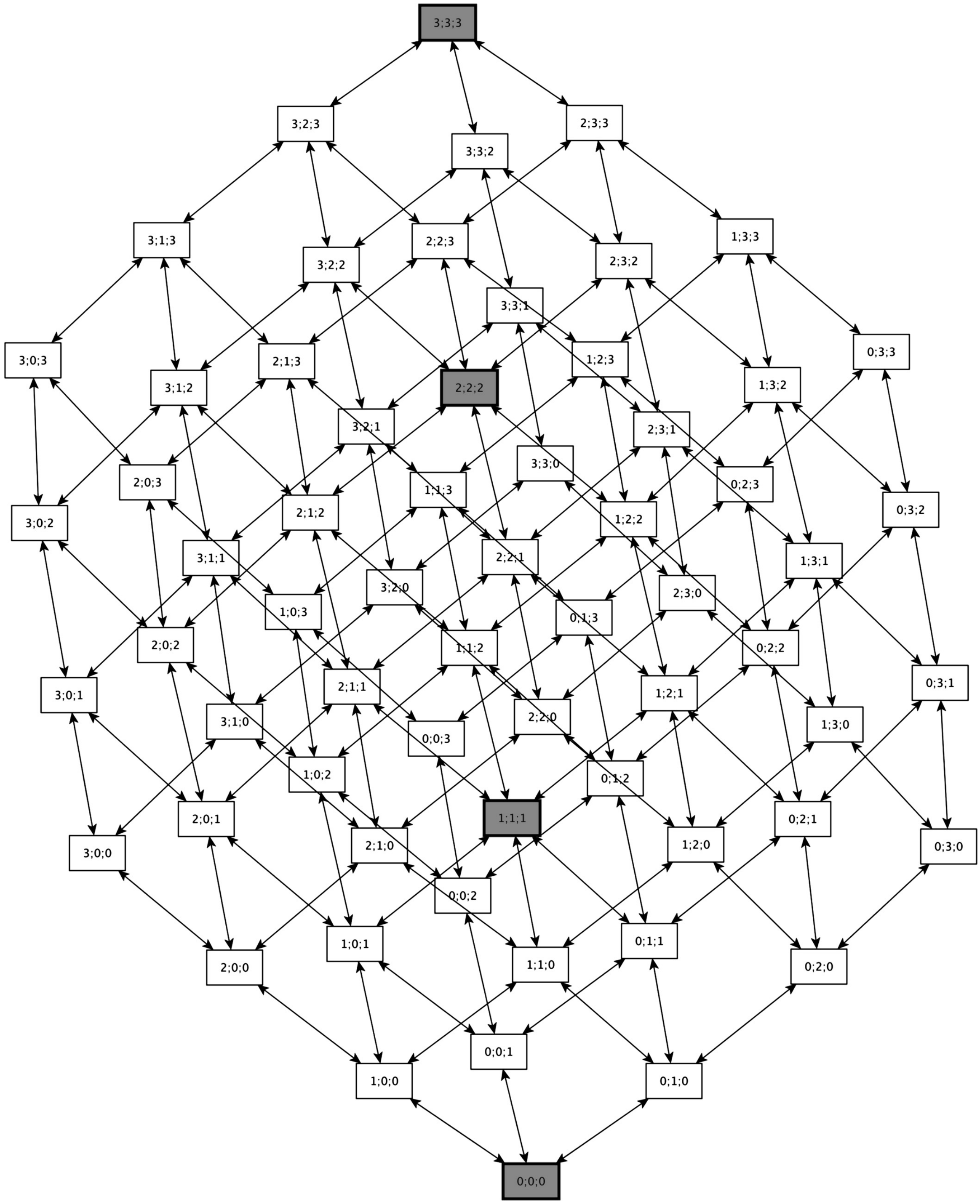

Figure A1. State-transition diagram of the MicroFIN task Fish-o-mat. Each rectangle represents one state of the task, with the numbers indicating the input levels for all three input variables (see Figure 1). Arrows represent transitions between states due to changes in input variable values. Gray states in the state-transition diagram are indicating states with equilibrium in input values, and hence, fish in the aquarium.

Table A1. Standardized trait and method factor loadings, consistency, and method specificity of MicroFIN and MicroDYN items

Instrument	Task	Item	CPS dimensions: Knowledge acquisition	CPS dimensions: Knowledge application	Method factors MicroFIN: Knowledge acquisition	Method factors MicroFIN: Knowledge application	Consistency	Method specificity
MicroFIN	Task 1: Concert-o-maton	Item 1	.502	.358	.572	.557	.435	.565
		Item 2					.292	.708
	Task 2: Plan-o-maton	Item 1	.510	.339	.351	.502	.678	.321
		Item 2					.313	.687
	Task 3: Plant-o-maton	Item 1	.499	.355	.458	.606	.543	.457
		Item 2					.255	.745
	Task 4: Fish-o-maton	Item 1	.511	.430	.442	.582	.572	.428
		Item 2					.353	.647
	Task 5: Flooz-o-maton	Item 1	.332	.308	.374	.439	.441	.559
		Item 2					.330	.670
MicroDYN	Task 1: Cat	Item 1	.842	.623			1	0
		Item 2					1	0
	Task 2: Moped	Item 1	.777	.817			1	0
		Item 2					1	0
	Task 3: Game night	Item 1	.810	.898			1	0
		Item 2					1	0
	Task 4: Perfume	Item 1	.897	.910			1	0
		Item 2					1	0
	Task 5: Gardening	Item 1	.896	.804			1	0
		Item 2					1	0
	Task 6: Handball team	Item 1	.729	.525			1	0
		Item 2					1	0
	Task 7: Spaceship	Item 1	.590	.315			1	0
		Item 2					1	0
	Task 8: First aid	Item 1	.750	.570			1	0
		Item 2					1	0

Notes. All loadings reported in the table are significant on a $p < .01$ level. The differences in loadings between MicroFIN and MicroDYN on the latent dimensions of CPS reverse when the reference method is changed from MicroDYN to MicroFIN.

Table A2. Linear structural equations, system size, and type of effects for the MicroDYN tasks

Task	Linear structural equations	System size	Effects
Task 1: Cat	$X_t + 1 = 1*X_t + 0*A_t + 2*B_t$	2×2 – System	Only effects of inputs
	$Y_t + 1 = 1*Y_t + 0*A_t + 2*B_t$		
Task 2: Moped	$X_t + 1 = 1*X_t + 2*A_t + 2*B_t + 0*C_t$	2×3 – System	Only effects of inputs
	$Y_t + 1 = 1*Y_t + 0*A_t + 0*B_t + 2*C_t$		
Task 3: Game night	$X_t + 1 = 1*X_t + 0*A_t + 2*B_t + 0*C_t$	3×3 – System	Only effects of inputs
	$Y_t + 1 = 1*Y_t + 2*A_t + 0*B_t + 0*C_t$		
	$Z_t + 1 = 1*Z_t + 0*A_t + 0*B_t + 2*C_t$		
Task 4: Perfume	$X_t + 1 = 1*X_t + 2*A_t + 0*B_t + 0*C_t$	3×3 – System	Only effects of inputs
	$Y_t + 1 = 1*Y_t + 0*A_t + 2*B_t + 2*C_t$		
	$Z_t + 1 = 1*Z_t + 0*A_t + 0*B_t + 2*C_t$		
Task 5: Gardening	$X_t + 1 = 1*X_t + 2*A_t + 2*B_t + 0*C_t$	3×3 – System	Only effects of inputs
	$Y_t + 1 = 1*Y_t + 0*A_t + 2*B_t + 0*C_t$		
	$Z_t + 1 = 1*Z_t + 0*A_t + 0*B_t + 2*C_t$		
Task 6: Handball team	$X_t + 1 = 1.33*X_t + 2*A_t + 0*B_t + 0*C_t$	2×3 – System	Effects of inputs and outputs
	$Y_t + 1 = 1*Y_t + 0*A_t + 0*B_t + 2*C_t$		
Task 7: Spaceship	$X_t + 1 = 1*X_t + 0*A_t + 0*B_t + 0*C_t$	3×3 – System	Effects of inputs and outputs
	$Y_t + 1 = 1.33*Y_t + 2*A_t + 2*B_t + 0*C_t$		
	$Z_t + 1 = 1*Z_t + 0*A_t + 0*B_t + 2*C_t$		
Task 8: First aid	$X_t + 1 = 1*X_t + 2*A_t + 0*B_t + 0*C_t$	3×3 – System	Effects of inputs and outputs
	$Y_t + 1 = 1*Y_t + 2*A_t + 0*B_t + 0*C_t$		
	$Z_t + 1 = 1.33*Z_t + 0*A_t + 0*B_t + 2*C_t$		

Notes. Features of the MicroDYN tasks. Linear Structural Equations: Values of the output variables (X, Y, Z) at time $t + 1$ depending on input and output (A, B, C) variables at time t. System size: Number of input and output variables. Effects: Only effects of input variables on output variables or effects of both, input and output variables on output variables (e.g., including Eigendynamics).

Detailed Description of the MicroFIN Tasks

Task 1: Concert-o-maton

The task "Concert-o-maton" features combinations of the input variables "music group" (two qualitatively different options) and "stage" (two options). Both can be varied independently. Additionally, the factor "admission fee" can be varied (three ordinal options).

Based on the combination of "music group" and "stage," the influence of "admission fee" on the number of visitors (outcome variable, three ordinal values) is varied. One option of both input variables is matched to another one, leading to variations in consequence to the height of the admission fee for a specific combination of "music group" and "stage." The effect of admission fee is reversed for a second combination of "music group" and "stage," while other combinations result in no influence of admission fee on the output variable at all.

Task 2: Plan-o-maton

Four input values (depictions of different buildings) are presented in a 2×2 matrix format. The position of the values can be exchanged on a bilateral level (e.g., changing the position of the picture from top left to top right, and vice versa) by pressing a button. Output values (four variables, each with two options) are presented between the pictures, their value based on the combination of the pictures (two pairs of the buildings can be matched). The combination of pictures leads to output values independently of the place the pictures are shown, just based on the combination.

Task 3: Plant-o-maton

Three input variables (four ordinal values each) are related to one output variable (three ordinal values). One of the input variables features a threshold value, after which the direction of influence on the output variable is changed from positive to negative. The other two variables are related to the output variable in a continuous positive and negative way.

Task 4: Fish-o-maton

Three input variables (four ordinal values each) can be manipulated, leading to different values in one output variable (an aquarium, five values). Variations in the output variable occur, if the input variables are brought to equivalent values (e.g., all input variable on medium setting leading to the display of a medium amount of fish). See also the article for a detailed description.

Task 5: Flooz-o-maton

Combinations of three different inputs with three values each (ordinal) have to be explored to create a cocktail. Three types of cocktails with different features (e.g., the amount of sugar) can be created, while some input combinations lead to no cocktail at all (i.e., threshold values). The microworld features a start button.

Intelligence 54 (2016) 55–69

Contents lists available at ScienceDirect

Intelligence

Construct validity of complex problem solving: A comprehensive view on different facets of intelligence and school grades☆,☆☆

André Kretzschmar [a,*], Jonas C. Neubert [a,b], Sascha Wüstenberg [a], Samuel Greiff [a]

[a] University of Luxembourg, 11, Porte des Sciences, 4366 Esch-Belval, Luxembourg
[b] Brandenburg University of Technology Cottbus-Senftenberg, Erich-Weinert-Straße 1, 03046 Cottbus, Germany

ARTICLE INFO

Article history:
Received 12 June 2015
Received in revised form 21 October 2015
Accepted 18 November 2015
Available online 4 December 2015

Keywords:
Complex problem solving
Intelligence
School grades
Nested factor model
Validity

ABSTRACT

Although complex problem solving (CPS) has attracted increasing amounts of attention in recent years (e.g., in the PISA study), the role of CPS in the nomological network of intelligence is controversial. The question of whether CPS is a distinct construct is as old as CPS research itself, but previous studies have had specific shortcomings when addressing the question of whether CPS is a separable or independent construct. The aim of the present study was, therefore, to combine the advantages of previous studies to facilitate a less biased view of the relation between CPS and established intelligence constructs. A sample of 227 German university students worked on a comprehensive measure of intelligence (Berlin Intelligence Structure test) and two CPS assessment tools (MicroDYN and MicroFIN). Furthermore, final school grades (GPA) served as an external criterion. We applied confirmatory factor analyses and structural equation modeling to investigate the relation between CPS and established intelligence constructs on the basis of different psychometric approaches (i.e., first-order model, nested factor model). Moreover, we examined the incremental validity of CPS in explaining GPA beyond established intelligence constructs. Results indicate that CPS represents unique variance that is not accounted for by established intelligence constructs. The incremental validity of CPS was found only when a commonly used narrow operationalization of intelligence was applied (i.e., figural reasoning) but not when a broad operationalization was applied.

In spite of the increasing popularity of complex problem solving (CPS),[1] especially in the educational sciences and international large-scale assessments such as the Programme for International Student Assessment (PISA; OECD, 2014), the status of CPS in the nomological network of intelligence is still controversial. Problem solving, in general, is seen as an essential part of intelligence (Gottfredson, 1997). However, there is a long-standing debate about whether CPS is just a new label for established constructs such as reasoning (fluid intelligence; e.g., Kröner, Plass, & Leutner, 2005; Süß, 1996) or a distinct cognitive construct not yet covered by established intelligence theories (e.g., Greiff, Wüstenberg, et al., 2013; Wüstenberg, Greiff, & Funke, 2012).

CPS describes the ability to solve unknown problem situations that are intransparent, dynamic, and interactive (e.g., Dörner, Kreuzig, Reither, & Stäudel, 1983, Frensch & Funke, 1995). This means, for instance, that relevant information needed to solve the problem is hidden from the outset (e.g., a new technical device without a manual). In order to solve the complex problem situation, the problem solver therefore needs to actively explore the problem situation to acquire knowledge (e.g., the functionality of controls). In a subsequent step, he or she can then use the acquired information to actually solve the problem (i.e., apply knowledge; Fischer, Greiff, & Funke, 2012). Accordingly, Wüstenberg et al. (2012) and Greiff, Fischer, Stadler, and Wüstenberg (2014) stated that the cognitive requirements associated with these dynamic interactions make CPS a separable construct as opposed to, for example, reasoning, which is usually measured with static tasks (e.g., all information needed to solve the problem is present and, thus, no new knowledge has to be acquired by interacting with the problem at hand).

On the other hand, the status of CPS in the nomological network of intelligence can also be viewed from a different angle (e.g., Kersting, 2001, Kröner et al., 2005, Süß, 1996). From this perspective, CPS is basically understood as a new label for or a conglomerate of already established

☆ We are grateful to the TBA group at DIPF (http://tba.dipf.de) for providing the authoring tool CBA Item Builder and technical support. We would also like to thank the various student assistants who assisted with the data collection and the data preparation and Philipp Sonnleitner for his comments on an earlier version of the manuscript. The data collection was part of the first author's dissertation.

☆☆ This research was partly funded by a grant from the Fonds National de la Recherche Luxembourg (ATTRACT “ASKI21”).

* Corresponding author at: ECCS Unit, University of Luxembourg, 11, Porte des Sciences, 4366 Esch-Belval, Luxembourg.

E-mail address: kretzsch.andre@gmail.com (A. Kretzschmar).

[1] Although we use the term *complex problem solving*, there are several synonyms in the literature, for example, interactive problem solving (e.g., Fischer et al., 2015), dynamic problem solving (e.g., Greiff et al., 2012), creative problem solving (e.g., OECD, 2014), or even dynamic decision making (e.g., Gonzalez, Thomas, & Vanyukov, 2005).

http://dx.doi.org/10.1016/j.intell.2015.11.004
0160-2896/

cognitive constructs. Reasoning, defined as "the use of deliberate and controlled mental operations to solve novel problems that cannot be performed automatically" (McGrew, 2009, p. 5), is thereby seen as a major cognitive ability that already includes the cognitive processes necessary to solve complex problems, including requirements such as the need to actively acquire knowledge. Consequently, some argue that the nomological network of intelligence does not need the distinct cognitive construct of CPS.

These two different perspectives (i.e., the arguments for a distinct CPS construct vs. the redundancy of CPS) have resulted in a controversial discussion that has existed since the beginning of CPS research (e.g., Dörner et al., 1983, Funke, 1999, Süß, 1996, Wüstenberg et al., 2012). Although both perspectives appear reasonable, previous studies from each perspective have their drawbacks. For example, the generalizability of previous findings is limited by psychometrically suboptimal CPS assessment tools, restricted operationalizations of intelligence, a lack of analyses on relations between CPS and external criteria, and a focus on specific psychometric approaches.

The purpose of the present study was to overcome these limitations and, thus, to shed further light on the issue of a distinct CPS construct, both theoretically and empirically. To do so, we address the origins of both perspectives and their empirical findings in the next section before presenting our empirical investigation.

1. Complex problem solving and intelligence: two perspectives

1.1. Redundancy perspective: complex problem solving as intelligence

From a theoretical point of view, it can be argued that there is a substantial overlap between CPS and established constructs of intelligence, in particular, reasoning. Mental operations such as drawing inferences, generating and testing hypotheses, identifying relations, comprehending implications, problem solving, extrapolating, and transforming information are seen as the core processes of reasoning (McGrew, 2009). At the same time, these operations closely correspond with the main mental operations applied in CPS (see Fischer et al., 2012; Greiff, Fischer, et al., 2014). Accordingly, Süß (1996, 1999) explicated that primarily processes of inductive and deductive reasoning are necessary to solve complex problems (e.g., detecting relations between a set of variables in a complex and dynamically changing system). However, the overlap between intelligence constructs and CPS is not limited to aspects of reasoning. Süß (1996) mentioned that additional intelligence constructs such as mental speed and crystallized intelligence (i.e., general and domain-specific knowledge) might also be involved if time constraints exist or if domain-specific problems need to be solved. In summary, CPS could be seen as a new (but redundant) label for established intelligence constructs or a conglomerate of them but not as a new construct that justifies the extension of current theories of intelligence.

This view has been empirically underpinned by several studies. For example, Süß (1996) demonstrated manifest correlations between a comprehensive operationalization of intelligence (i.e., reasoning, mental speed, memory, and creativity on the Berlin Intelligence Structure Test, BIS; Jäger, Süß, & Beauducel, 1997; Süß & Beauducel, 2015; as well as several tests of crystallized intelligence) and CPS (assessment tool: Tailorshop; Putz-Osterloh, 1981) up to $r = .65$, an overall amount of variance explained in CPS in a multiple regression with different facets of intelligence of up to 51%, and no significant correlation between two measures of CPS at different points in time when controlling for these facets of intelligence. Other studies (Süß, 1999; Wittmann & Süß, 1999) found manifest correlations between established intelligence constructs (BIS test) and several instruments targeting CPS (assessment tools: LEARN!, Milling, 1996; Tailorshop, Putz-Osterloh, 1981; PowerPlant, Wallach, 1997) of up to $r = .56$ and about 32% variance explained in CPS. Again, the correlations between CPS measurement instruments were nonsignificant when the broad operationalization of intelligence using the BIS test was controlled for. Other studies from this perspective did not even distinguish between CPS and intelligence but rather used CPS assessment tools as interactive measures of reasoning. For instance, Kröner et al. (2005) interpreted manifest correlations between CPS (assessment tool: MultiFlux; Kröner, 2001) and reasoning (BIS subscale) of $r = .67$ as evidence for convergent validity between two different intelligence measures—one using a classical paper-pencil test and one using a computer-based dynamic assessment environment.

In summary, studies from the redundancy perspective have reported (mainly manifest) high correlations between established intelligence constructs and CPS. In fact, these correlations were described as being as high as "[…] one would expect from a typical correlation between conventional intelligence tests" (Kröner et al., 2005, p. 365). In addition, it was argued that systematic variance in CPS could be fully explained with established intelligence constructs (Kersting, 2001; Süß, 1996, 1999). Both criteria (i.e., the high correlation between intelligence measures and CPS and the absence of systematic CPS variance) led to the conclusion that there was no evidence for a specific CPS construct.

1.2. Distinctness perspective: complex problem solving as a separate construct

Acknowledging the studies mentioned above, proponents of the distinctness perspective confirmed an overlap between CPS and established intelligence constructs but emphasized the unique cognitive requirements of CPS. According to this perspective, solving a complex problem requires the problem solver to deal with a lack of information at the outset, actively generate information, deal with dynamic interactions, and use procedural knowledge (Greiff, Fischer, et al., 2014a; Putz-Osterloh, 1981). Thus, more complex cognitions must be involved in CPS to handle the dynamic interactions in complex problems—in particular in comparison with simple cognitions (e.g., processing capacity, mental speed; see Funke, 2010), which would be fairly well-covered by traditional intelligence tests such as Raven's Advanced Progressive Matrices (APM; Raven, 1958). In other words, the argument is that established constructs of intelligence might not be sufficient to cover the mental processes involved in CPS. Hence, CPS is seen as a distinct construct: located in the nomological network of intelligence but separable from established constructs such as reasoning. It is important to note that agreement has not been achieved about exactly where to locate CPS in concurrent theories of intelligence (see Danner, Hagemann, Schankin, Hager, & Funke, 2011; Wüstenberg et al., 2012).

The distinctness perspective has also been empirically supported by a number of studies. For example, Wüstenberg et al. (2012) reported a latent correlation between figural reasoning (APM) and CPS (assessment tool: MicroDYN; Greiff, Wüstenberg, & Funke, 2012) of up to $r = .63$ and a proportion of variance explained in CPS of up to 39%. Furthermore, a significant and strong correlation between tasks targeting CPS and even between two CPS subprocesses (i.e., knowledge acquisition and knowledge application; Fischer et al., 2012) were also found when controlling for figural reasoning ability. In contrast to previous studies, Wüstenberg et al. (2012) additionally reported incremental predictive validity[2] for CPS. CPS explained incremental variability in final school grades (grade point average; GPA) beyond figural reasoning (6% additional explained variance), indicating an incremental utility of CPS beyond an established intelligence construct. Other studies have replicated these findings several times (e.g., Greiff, Fischer, et al., 2013; Greiff, Wüstenberg, et al., 2013). In a different study, Danner, Hagemann, Holt, et al. (2011) reported manifest correlations between figural reasoning (APM) and two instruments targeting CPS (assessment tools: Tailorshop, Putz-Osterloh, 1981; HEIFI, Wirth & Funke, 2005) of up to $r = .55$ and interpreted this finding as indicative of separable

[2] The term *predictive* intuitively belongs to longitudinal studies, but it is often used in cross-sectional studies as well (e.g., Wüstenberg et al., 2012). Following this practice, it is also used here for statistically explaining variance in criteria.

constructs. In addition, the manifest correlation between the two CPS measures remained significant ($r = .20$) when controlling for figural reasoning. Furthermore, an incremental prediction of supervisory ratings with one of the CPS assessment tools beyond figural reasoning was established ($r = .22$). Finally, Sonnleitner, Keller, Martin, and Brunner (2013) reported latent correlations between reasoning (three tasks of IST-2000R; Amthauer, Brocke, Liepmann, & Beauducel, 2001) and CPS (assessment tool: Genetics Lab; Sonnleitner et al., 2012) of up to $r =$ .62. In addition, a specific CPS factor (i.e., independent of a *g*-factor) was established in a nested factor model. However, the specific CPS factor hardly predicted school grades beyond reasoning when a nested factor measurement model was applied.

In summary, studies from the distinctness perspective have reported (mainly latent) correlations, suggesting "[…] that reasoning ability and CPS may represent distinct cognitive abilities […]" (Sonnleitner et al., 2013, p. 300). In addition, systematic variance in CPS was not fully explained by the employed intelligence assessments, and CPS often incrementally explained variance in external criteria beyond reasoning. These findings led to the conclusion that there was substantial evidence for a distinct CPS construct in the nomological network of intelligence.

1.3. Similarities and differences between the two perspectives

From a methodological point of view, the empirical studies from the two perspectives differ in a variety of features such as CPS assessment tools, operationalizations of intelligence, the examination of relations to external criteria, and applied psychometric approaches (see Table 1 for an overview). A direct comparison between the findings from the two perspectives is therefore challenging. However, a closer look reveals some characteristics of the studies that might explain the different findings and interpretations from the redundancy and distinctness perspectives.

1.3.1. CPS assessment tools

In recent years, the assessment of CPS has evolved, and different assessment tools have been used in the respective studies from the two perspectives. Semantically rich and comprehensive microworlds were predominantly applied in the beginning of CPS research and can be found in particular in studies from the redundancy perspective (e.g., Kersting, 2001, Süß, 1996, 1999). In such microworlds (e.g., Tailorshop), the complex problem presented to test takers is embedded in a real-world context (e.g., managing a company) to ensure high *ecological validity* (Funke, 2001; Süß, 1999). As a consequence of this embedding, such semantically rich microworlds rely heavily on the problem solver's prior domain-specific knowledge (Funke, 1992b; Süß, 1996). For instance, prior knowledge such as understanding the principles of marketing is relevant for managing the Tailorshop simulation. Consequently, participants' CPS performance in such microworlds is confounded with prior domain-specific knowledge (i.e., crystallized intelligence), which, in turn, might interfere with a reliable and valid assessment of a domain-general CPS ability (Greiff, 2012; Süß, 1996).

In an alternative approach (Funke, 1992a), the utility of prior knowledge is therefore strongly reduced by avoiding deep semantic meaning when labeling input and output elements (e.g., by using abstract labels such as "Control A"; see also Beckmann & Goode, 2014). In this way, CPS can be measured with only a minimal influence of prior knowledge in a semantically poor approach. Another shortcoming of earlier CPS assessments, the rather low reliability of the measurement due to long testing times and a small number of indicators, was overcome by developing the multiple-task approach (Greiff, 2012). Instead of applying a comprehensive one-task assessment tool for at least 1 h, multiple small and semantically poor complex systems (time-on-task for each task of about 5 to 10 min) are used in assessment instruments such as the Genetics Lab (Sonnleitner et al., 2012) or MicroDYN (Greiff et al., 2012). These assessment instruments exhibit good psychometric qualities in terms of reliability while keeping the testing time within practical limits, thereby improving upon the previously employed microworlds that required immense amounts of time to yield acceptable reliability (e.g., Süß, 1996, Wittmann & Süß, 1999).

The differences in the findings between the redundancy and distinctness perspectives may be attributed to the use of different CPS assessment tools. Recent studies, and in particular those from the distinctness perspective, have mainly been conducted with the newly developed assessment tools (e.g., Neubert, Kretzschmar, Wüstenberg, & Greiff, 2015a, Sonnleitner et al., 2013, Wüstenberg et al., 2012). The criteria used to evaluate the relation between CPS and established intelligence constructs could therefore have been biased by the suboptimal psychometric features of early assessment tools.

1.3.2. Operationalization of intelligence

In addition to the selection of appropriate assessment tools for targeting CPS, it is also important to consider the operationalization of intelligence when studying the relation between established intelligence constructs and CPS. In the redundancy perspective, CPS has often been examined with the comprehensive Berlin Intelligence Structure (BIS) test (Jäger et al., 1997) and the corresponding theoretical approach of the BIS model (Jäger, 1984; for a description in English, see Süß & Beauducel, 2015). In this model, intelligence is organized in a faceted structure (see Fig. 1) featuring operation factors (i.e., reasoning, mental speed, creativity, memory), content factors (i.e., figural, verbal, numerical), and a general intelligence factor (*g*). The construction of the BIS test was based on a nearly representative sample of all intelligence test tasks documented in the literature at that time. Therefore, the BIS test can be considered a highly comprehensive and construct-valid operationalization of intelligence (Süß & Beauducel, 2015).

Studies from the distinctness perspective, however, have usually applied narrow instead of broad and thus more time-consuming operationalizations of intelligence (e.g., relying on figural reasoning such as indicated by the APM; see Danner, Hagemann, Holt, et al., 2011; Greiff, Wüstenberg, et al., 2013; Neubert et al., 2015a; Wüstenberg et al., 2012). Although specific operationalizations such as the APM or the Culture Fair Test 20-R (CFT; Weiß, 2006) are assumed to be good indicators of *g*, the construct validity of *g* is reduced if only such specific operationalizations are used (Shadish, Cook, & Campbell, 2001; Süß & Beauducel, 2011). As seen in Fig. 1, the APM and CFT as two tests that have been heavily utilized in CPS studies from the distinctness perspective are located in one specific cell of the BIS model (i.e., figural reasoning), contradicting the requirements for a construct-valid operationalization of intelligence (see Gignac, 2015; Jensen & Wang, 1994; Reeve & Blacksmith, 2009).

Table 1
Key characteristics of typical studies from the redundancy and distinctness perspectives and the present study.

Study characteristics	Redundancy perspective	Distinctness perspective	Present study
CPS assessment tools	Semantically rich microworlds	Semantically abstract multiple complex systems	Semantically abstract multiple complex systems
Operationalization of intelligence	Broad	Narrow	Broad and narrow
Incremental predictive validity	Rarely examined	Almost always examined	Examined
Measurement models	Almost always first-order factor model	Almost always first-order factor model	First-order factor model and nested-factor model

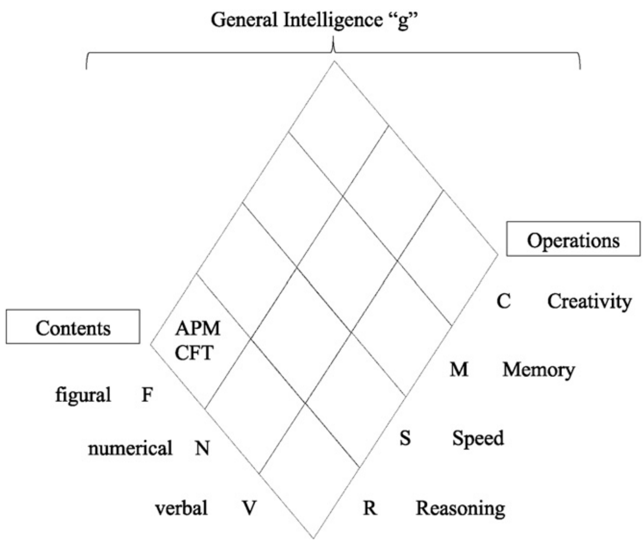

Fig. 1. The Berlin Intelligence Structure (BIS) Model. To demonstrate the narrow operationalization of intelligence, the Advanced Progressive Matrices (APM) and the Culture Fair Test 20-R (CFT) are located in one specific cell of the BIS model.

The possible consequences of such narrow operationalizations were impressively demonstrated by Beauducel, Liepmann, Felfe, and Nettelnstroth (2007) in a different context: Whereas a broad operationalization of intelligence led to significant correlations between intelligence constructs and personality dimensions, figural reasoning as a widely used narrow operationalization did not result in any significant correlations with personality measures. Thus, the generalizability of findings seems to be highly dependent on the breadth of the operationalization of intelligence. Considering this dependency, the rather moderate correlation between CPS and established intelligence constructs in the studies from the distinctness perspective may have been caused by the application of specific and narrow operationalizations of intelligence.

1.3.3. Incremental predictive validity

Although evidence of convergent and discriminant validity is important when examining different constructs, incremental predictive validity must also be examined (Kersting, 2001; Süß, 1999). In the context of CPS, this would mean that CPS must be able to explain variance in external criteria (e.g., school grades) that goes above and beyond other crucial predictors such as established intelligence constructs (see Kuncel, Hezlett, & Ones, 2004). Even if the relation between two constructs is rather high, the presence of incremental predictive validity can illustrate the utility of an additional construct (for an example in the context of intelligence and working memory, see Lu, Weber, Spinath, & Shi, 2011).

In CPS research, many studies from the distinctness perspective have explored the incremental validity of CPS (and almost always found evidence for it, e.g., Greiff, Fischer, et al., 2013; Greiff, Kretzschmar, Müller, Spinath, & Martin, 2014; Kretzschmar, Neubert, & Greiff, 2014; Wüstenberg et al., 2012), whereas studies from the redundancy perspective have rarely examined the relation of CPS to external criteria (for an exception, see Kersting, 2001). Therefore, it might be possible that the redundancy perspective's arguments against a CPS construct have been premature as the incremental predictive validity of CPS has not been considered in those studies.

1.3.4. Measurement models

As the reader may have already noticed from the descriptions of previous empirical findings, different statistical approaches were used in different studies (e.g., manifest correlations vs. structural equation modeling). Considering the advancements in statistical analyses (e.g., Nachtigall, Kroehne, Funke, & Steyer, 2003), latent analyses as applied in recent studies from the distinctness perspective seem to be more suitable for examining the relation between CPS and intelligence on a construct level as they were able to reduce the influence of measurement error.

By applying latent analyses, Sonnleitner et al. (2013) emphasized the impact of different measurement models in CPS research. Usually, a first-order factor model approach has been applied to examine the relation between CPS and established intelligence constructs (e.g., Kröner et al., 2005, Wüstenberg et al., 2012). With this approach (see Fig. 2a), individual performance on one instrument is ascribed only to the corresponding construct (i.e., only to a CPS factor or to an intelligence factor, respectively). To examine the relation between CPS and established intelligence constructs, correlational analyses between the two latent factors have often then been applied.

However, as Brunner (2008) noted, the first-order factor model approach might neglect the impact of *g* with crucial consequences. From a theoretical point of view, it is widely accepted that there is a *g*-factor underlying all cognitive abilities (see e.g., the BIS Model; Jäger, 1984; or CHC Theory; McGrew, 2009). Following this line of argumentation, the CPS factor in a first-order factor model would consist not only of unique CPS variance but also of *g*-factor variance. Thus, the relation between CPS and, for example, external criteria might be mainly caused by the *g*-factor variance rather than CPS variance. However, the presence of a *g*-factor can be directly accounted for on the level of measurement models by using a nested factor model approach (Brunner, Nagy, & Wilhelm, 2012; Reise, 2012; see Fig. 2b). In this approach, specific factors account for only unique variance that is independent of the *g*-factor. Therefore, the specific factors can be used to examine "pure" specific ability (Murray & Johnson, 2013)—an issue that is particularly important when examining incremental predictive validity (see above).[3]

Notably, almost all of the studies from both the redundancy and distinctness perspectives have used a first-order factor model approach. As an exception, Sonnleitner et al. (2013) additionally used the nested factor measurement model. We will illustrate the potential consequences in the context of the incremental predictive validity of CPS. Although the correlational pattern for a first-order factor model for CPS and GPA was higher in Sonnleitner et al.'s study (2013; $R^2 = 14$–16%) than in previous studies (e.g., Wüstenberg et al., 2012; $R^2 = 10\%$), Sonnleitner et al. found only a marginal incremental predictive power of CPS when using a nested factor measurement model. Hence, it might be possible that previous studies using the first-order factor model approach overestimated the distinctness of a potential CPS construct due to their use of first-order factor models.

1.4. The present study

The goal of the present study was to examine the role of CPS in the nomological network of intelligence. To this end, we combined the crucial advantages of studies from the redundancy and distinctness perspectives, as outlined above, into one empirical study. More specifically, we used psychometrically sound CPS assessment tools from the multiple-task approach (i.e., MicroDYN and MicroFIN) as applied in the studies from the distinctness perspective. Furthermore, we used a comprehensive operationalization of intelligence (i.e., BIS test and crystallized intelligence test) as done in the studies from the redundancy perspective. To compare our results with those from the distinctness-perspective

[3] Please note that several studies (e.g., Greiff, Fischer, et al., 2013; Wüstenberg et al., 2012) used the residual variance of CPS in a latent stepwise regression in order to overcome the limitations of the first-order factor model approach when studying the incremental predictive validity of CPS. However, the applied latent stepwise regression approach and the nested factor model approach are not equivalent.

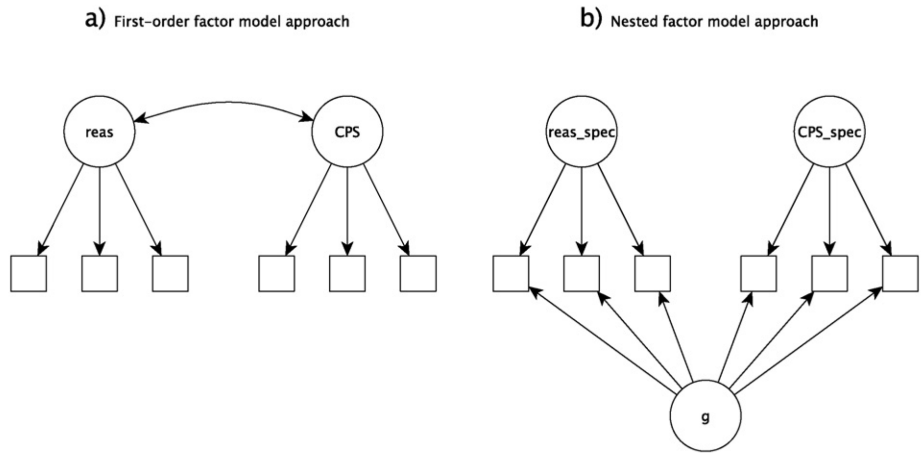

Fig. 2. The first-order factor model and the nested factor model approach with regard to reasoning and CPS. reas: reasoning ability; CPS: complex problem solving ability; reas_spec: specific reasoning ability; CPS_spec: specific complex problem solving ability; g: general factor of intelligence (*g*-factor).

studies, we additionally included analyses that were based on a narrow operationalization of intelligence (i.e., figural reasoning as a common narrow operationalization). Moreover, we examined the incremental predictive validity of CPS with regard to a widely used external criterion (i.e., school grades). For all analyses, we used two different measurement model approaches (i.e., first-order factor model vs. nested factor model) to investigate their impact with regard to the following research issues (see Table 1 for a summary).

1.5. Research issue 1: relation between CPS and established constructs of intelligence

Our first research issue comprises the relations between CPS and established intelligence constructs. To ensure the comparability of our findings with previous studies, we focused on two methods of analysis commonly applied in previous CPS research: the amount of variance explained when CPS is statistically predicted by established intelligence constructs using the first-order factor model approach (e.g., Süß, 1996, Wüstenberg et al., 2012) and the application of confirmatory factor analyses to investigate CPS using a nested factor model approach (Sonnleitner et al., 2013).

Our first method of analysis using the first-order factor model approach (see Fig. 2a) identified the amount of variance explained in CPS performance statistically predicted by established intelligence constructs. If a substantial amount of CPS variance were to remain unexplained in these analyses, the notion of a separable construct of CPS would receive further support. However, we were unable to identify a consistent criterion for representing a "substantial amount" of unexplained variance. A review of the CPS literature revealed a proportion of unexplained variance in CPS of 58% as supportive of the redundancy perspective (i.e., speaking for a common construct; Kröner et al., 2005; Süß, 1996) but also 60% as favoring the distinctness perspective (Wüstenberg et al., 2012). It is interesting that proponents of both perspectives used almost the same criterion of roughly 60% of the variance in CPS left unexplained as supporting their view of a redundant or distinct CPS construct, respectively. As a consequence, we therefore decided not to define a specific threshold for a distinct CPS construct but rather to compare our findings with previous studies from both perspectives. As we combined the advantages of different previous studies (e.g., reliable assessment of CPS, broad operationalization of intelligence) to facilitate a less biased view of the relations between CPS and established intelligence constructs, we expected a smaller proportion of unexplained variance in CPS than in previous research. Hence, Hypothesis 1.1 predicted that less than 60% of the variance in CPS would remain unexplained when predicting CPS with established intelligence constructs.

The second method of analysis refers to the unique and systematic variance in CPS in a measurement model that combines CPS and several intelligence constructs. That is, if CPS is distinct from established intelligence constructs, unique and systematic variance should be demonstrable even when established intelligence constructs are controlled for. To do so, we used confirmatory factor analyses to apply a nested factor model (see Fig. 2b). The model contained a *g*-factor and specific (i.e., orthogonal) factors for specific cognitive abilities (see Brunner et al., 2012; Reise, 2012). To test the assumption of a distinct CPS construct, we expected a nested factor model with a *g*-factor (i.e., common variance between established intelligence constructs and CPS), specific factors for established intelligence constructs (i.e., the unique variance of mental speed, memory, etc.), and a specific factor for CPS (i.e., the unique variance of CPS) to hold. Therefore, Hypothesis 1.2 predicted that if a nested factor model was used to model CPS and established intelligence constructs, a substantial specific CPS factor would emerge.

1.6. Research issue 2: incremental predictive validity of CPS

Incremental predictive validity can be taken as strong evidence of a distinct construct. Hence, studies supporting the distinctness perspective have usually examined the incremental predictive validity of CPS beyond established intelligence constructs in explaining variance in external criteria. With regard to the choice of an external criterion, CPS can be seen as providing an important cognitive foundation for academic achievement, and thus, school grades have commonly been used as external criteria in previous CPS research (see Greiff, Wüstenberg, et al., 2013; Kretzschmar et al., 2014; Sonnleitner et al., 2013; Wüstenberg et al., 2012). Following this approach, we statistically predicted GPA with CPS in addition to several intelligence constructs to investigate the incremental predictive validity of a potential CPS construct.

As outlined above, the measurement model (i.e., first-order model approach vs. nested factor model approach, see Fig. 2) might have a substantial impact on the incremental predictive validity of CPS. Therefore, we used both approaches to examine Research Issue 2. In the first-order factor model, we first predicted school grades with established intelligence constructs only. In a second step, CPS was added as an additional predictor of school grades. A significant amount of additional variance in school grades explained by CPS in the second model could then be interpreted as evidence for the incremental predictive validity of CPS.

Thus, Hypothesis 2.1 predicted that using the first-order factor model approach, CPS would be positively related to school grades and would incrementally explain variance in school grades beyond established intelligence constructs.

In the nested factor model approach, school grades were predicted by the *g*-factor, the specific factors of established intelligence constructs, and CPS. Testing the incremental predictive validity of CPS, we expected that the specific factor of CPS would significantly and positively predict school grades beyond the *g*-factor and the other specific factors. Therefore, Hypothesis 2.2 predicted that using the nested factor model approach, the specific CPS factor would be positively related to school grades and would incrementally explain variance in school grades beyond the *g*-factor and the other specific factors.

In summary, we used different methods and criteria to investigate the role of CPS in the nomological network of intelligence. More precisely, we applied a 2 × 2 × 2 research design with regard to (a) criteria to investigate a distinct construct (the relation between different established intelligence constructs and CPS vs. the incremental prediction of school grades by CPS beyond the intelligence constructs), (b) the impact of the breadth of the operationalization of intelligence (i.e., a broad operationalization with several intelligence constructs vs. a narrow operationalization with only figural reasoning), and (c) the influence of different measurement model approaches (i.e., first-order factor model vs. nested factor model) on our findings.

2. Method

2.1. Participants

The sample was part of a larger study that was conducted at the University of Heidelberg, Germany. The sample was comprised of $N = 227$ university students (73% female; age: $M = 22.88$, $SD = 4.27$) who volunteered to take the CPS assessment. Participants received course credit or 40€ for their participation.

2.2. Materials

2.2.1. Complex problem solving

CPS was assessed with two different computer-based assessment tools (MicroDYN and MicroFIN) to ensure generalizability beyond the specific operationalizations. Both measurements are based on a multiple-task approach and formal frameworks (e.g., Greiff et al., 2012, Neubert et al., 2015a). Furthermore, both cover the two core processes of CPS, namely, knowledge acquisition and knowledge application (see Fischer et al., 2012).

2.2.1.1. MicroDYN. The MicroDYN approach (Greiff et al., 2012) is based on the formal framework of linear structural equations (Funke, 1985). MicroDYN is a reliable (Cronbach's αs > .70; e.g., Wüstenberg et al., 2012) measurement tool that has been utilized in large-scale educational assessments (e.g., OECD, 2014). Previous research has provided evidence for its construct (e.g., Greiff, Fischer, et al., 2013) and incremental predictive (e.g., Greiff, Wüstenberg, et al., 2013; Kretzschmar et al., 2014) validity. A typical MicroDYN task first asks participants to explore an unknown system to detect and note causal relations between several input and output variables (i.e., knowledge acquisition). Subsequently, they are asked to manipulate the system in order to reach a given goal (i.e., knowledge application). Apart from the introductory task, participants completed six MicroDYN tasks, resulting in a total processing time of about 40 min for MicroDYN (see Appendix Table A1 for formal task descriptions). Each task was scored according to Wüstenberg et al.'s (2012) procedure (i.e., each knowledge acquisition and knowledge application subtask was scored dichotomously).

2.2.1.2. MicroFIN. The MicroFIN approach (Neubert et al., 2015a) adapts the advantages of MicroDYN (i.e., especially high reliability due to multiple tasks) while simultaneously aiming for greater heterogeneity in CPS assessment. It is based on the formal framework of finite state automata (Buchner & Funke, 1993). Previous research (e.g., Neubert et al., 2015a) has indicated sufficient reliability (McDonald's ω ≥ .78) and an overlap with other instruments targeting CPS (i.e., correlation between MicroDYN and MicroFIN: r ≥ .56). As in MicroDYN, participants are first asked to explore an unknown system. Afterwards, questions about their acquired knowledge concerning the system are asked (i.e., knowledge acquisition). Subsequently, participants have to reach specific goals (i.e., knowledge acquisition). In this study, we used one introductory and six heterogeneous MicroFIN tasks (see Appendix B for the task descriptions) resulting in a total processing time of about 60 min. Each task was scored according to Neubert et al.'s (2015a) procedure (i.e., subtasks were scored dichotomously, and a sum score for each task was computed for knowledge acquisition and knowledge application, respectively).

2.2.2. Intelligence

Intelligence was assessed with the Berlin Intelligence Structure Test (BIS-4 Test; Jäger et al., 1997). The test is based on the Berlin Intelligence Structure Model by Jäger (1984). The BIS model describes a faceted structure of intelligence with four operation factors (i.e., reasoning, mental speed, memory, creativity), three content factors (i.e., figural, numerical, verbal), and a *g*-factor (for an English description of the model and the test, see Süß & Beauducel, 2015). The BIS test contains 45 tasks and takes a total of approximately 2.5 h. The test was administered and the scores were computed according to the test manual.[4] In line with our hypotheses, cognitive processes were of particular interest; hence, we included only the scores for the four operation factors (i.e., reasoning, mental speed, memory, and creativity) in our analyses and did not use the content factors.

General knowledge (in terms of crystallized intelligence; see Carroll, 1993; McGrew, 2009) is not part of the original BIS model or the BIS test (Beauducel & Kersting, 2002; Süß & Beauducel, 2011). However, crystallized intelligence is an integral facet of several intelligence models (e.g., CHC theory; McGrew, 2009). Therefore, we extended the measurement of intelligence by using the Bochumer Knowledge Test (BOWIT) as a measure of crystallized intelligence (Hossiep & Schulte, 2008) to ensure an even broader operationalization of intelligence. We used the short 45-question version, which covers the two domains social/society sciences and natural/technical sciences. The test takes about 20 min and was scored according to the manual.

2.2.3. School grades

As in previous CPS research (e.g., Fischer et al., 2015, Wüstenberg et al., 2012), academic achievement was measured with self-reported final school grade point average (GPA; for the validity of self-reported grades, see Bahrick, Hall, & Berger, 1996; Sparfeldt, Buch, Rost, & Lehmann, 2008). In German school systems, school grades range from 1 (*excellent*) to 6 (*insufficient*). For our analyses, school grades were reversed so that higher numbers reflected better performance.

2.3. Procedure

Testing was conducted in two successive sessions. The first session (approximately 2 h) included the assessment of CPS, school grades, and demographic data. The second session (approximately 3 h) was usually conducted within 1 week of the first session. In this second

[4] To score the creativity tasks, we used the scoring procedure for fluency (U mode; Jäger et al., 1997).

session, participants worked on the intelligence tests and additional questionnaires that were not part of this article.

2.4. Data analyses

We used the R software (version 3.0.2; R Core Team, 2013) with the packages lavaan (version 0.5–17.711; Rosseel, 2012) and psych (version 1.5.1; Revelle, 2015) as well as Statistics Calculators (version 3.0; Soper, 2015) for our analyses. The data and R syntax for the following analyses are publicly available via the Open Science Framework and can be accessed at https://osf.io/qf673/.

We examined the measurement models by computing confirmatory factor analyses (CFA). Established measurement models from previous research (i.e., MicroDYN and MicroFIN) as well as a measurement model that was in line with the BOWIT test description were first analyzed on the basis of single items. As all of the MicroDYN, MicroFIN, and BOWIT items were dichotomous, we used weighted least squares means and variance adjusted (WLSMV) estimation (see Moshagen & Musch, 2014). In a next step, parcel scores (i.e., mean scores of subsets of items) that were designed according to the item-to-construct principle (Little, Cunningham, Shahar, & Widaman, 2002) were created for these measurement models for further analyses. For the BIS test's measurement model, we calculated standard parcel scores on the basis of the theoretical assumptions behind the BIS model (see Jäger et al., 1997). We used parcels in order to increase the accuracy of the parameter estimates and to better capture the latent construct (Little, Rhemtulla, Gibson, & Schoemann, 2013). The measurement models were identified by fixing the variance of the latent factor to 1.00; all other model parameters were estimated freely if not stated otherwise. Building on the measurement models (based on parcels scores), we tested our hypotheses with structural equation modeling (SEM) with robust maximum likelihood (MLR) estimation. All reported coefficients for CFA and SEM were based on completely standardized solutions.

The evaluation of model fit was based on standard fit indices and the recommended cutoff values (Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003). In detail, we consulted the χ^2 goodness-of-fit statistic with the Yuan-Bentler correction, Comparative Fit Index (CFI > .95), Gamma Hat[5] (Gamma > .95), Standardized Root Mean Square Residual (SRMR < .08), Weighted Root Mean Square Residual (WRMR < .90), and Bayesian Information Criterion (BIC; lower values indicate better model fit). In order to compare different models (see Hypothesis 1.2), we used the Satorra-Bentler-scaled χ^2 difference test (i.e., nonsignificance indicates equal model fit) and differences between BIC values (ΔBIC > 10 indicates meaningful differences; Raftery, 1995). For Hypothesis 1.2, we calculated the reliability index omega specific (ω_s) for the specific factors in the nested factor model to evaluate whether a specific CPS factor provided reliable variance (see Reise, 2012; Reise, Bonifay, & Haviland, 2013). According to Reise et al.'s (2013) preliminary recommendations, $\omega_s > .50$ indicates sufficient and $\omega_s > .75$ good reliability for a specific factor (i.e., a specific factor provides reliable information that is unique from the *g*-factor).

A few participants took part in only the first but not in the second test session. Missing data on the intelligence test for 13% of the participants were due to this dropout. The percentage of missing data (e.g., due to software errors) was below 8% for all other measurements. We used the full information maximum likelihood (FIML) procedure to adjust for missing data. In general, tests of significance ($\alpha = .05$) were two-tailed except for the test of the direct hypothesis from Research Issue 2.

3. Results

3.1. Measurement models and descriptive statistics

First, we examined the measurement models for the different tests and assessment tools. The fit statistics and the range of factor loadings for each measurement model are reported in Table 2. All factor loadings were significant. Descriptive statistics for each measurement and correlations are presented in Table 3.

For MicroDYN, we applied a higher order factor model with a MicroDYN second-order factor at the top and two lower order factors for knowledge acquisition and knowledge application (see Greiff & Fischer, 2013; Kretzschmar et al., 2014). The factor loadings of the lower order factors on the second-order factor were constrained to be equal for identification purposes. The measurement model showed a good fit (see Table 2, M01). Therefore, in further analyses, a model with three parcels for knowledge acquisition and three parcels for knowledge application was applied and also showed a good fit (see Table 2, M02).

We applied the same higher order measurement model with two lower order factors for knowledge acquisition and knowledge application to MicroFIN (see Neubert et al., 2015a). However, the model did not demonstrate an acceptable fit (see Table 2, M03). Further modifications (e.g., one-dimensional model) did not improve the model fit sufficiently. Therefore, we examined a different measurement model in which a sum score combining knowledge acquisition and knowledge application was calculated for each MicroFIN task. The resulting six scores (i.e., one for each task) were then used to generate a one-dimensional model with a general latent MicroFIN factor. The one-dimensional model showed a good fit[6] (see Table 2, M04) and, thus, a measurement model with three parcels was applied (i.e., a just-identified measurement model).

Furthermore, we combined MicroDYN and MicroFIN into a general CPS factor to account for systematic CPS variance between different CPS assessment tools. That is, we tested a higher order factor model that incorporated the MicroDYN and MicroFIN measurement models. The factor loadings of the MicroDYN factor and the MicroFIN factor on the general CPS factor were constrained to be equal for identification purposes. The model showed a good fit (see Table 2, M05). All further analyses involving CPS were conducted with this general CPS factor model.

For the BIS test, we applied a first-order factor model with four correlated latent factors (i.e., reasoning, mental speed, memory, creativity). According to the BIS model and the recommended scoring procedure (i.e., suppressing unwanted error variance within a parcel, see Jäger et al., 1997), three to four theory-based parcels were used for each latent factor. The measurement model showed a very good fit (see Table 2, M06). To examine the narrower operationalization of intelligence, we also tested an additional measurement model with only the figural reasoning tasks from the BIS test. The one-dimensional model for the five tasks showed a very good fit (see Table 2, M07).

To our knowledge, no measurement model for the short version of the BOWIT has been published in previous research. Therefore, we applied a first-order factor model with two correlated latent factors (i.e., a social/society knowledge factor and a natural/technical knowledge factor) according to the theoretical assumptions of the test (Hossiep & Schulte, 2008) as well as a one-dimensional model (i.e., a general knowledge factor). Neither model showed an acceptable fit (see Table 2, M08 and M09). However, it was not our main interest to develop a completely new measurement model for a specific test,

[5] Please note that Gamma is closely related to the more frequently reported Root Mean Square Error of Approximation (RMSEA), but Gamma is less sensitive to different model types (Fan & Sivo, 2007).

[6] To ensure a unidimensional measurement model, we cross-validated the MicroFIN measurement model from the present study with Neubert et al.'s (2015a) data. Although somewhat different tasks were used in Neubert et al., the model fit was good: $\chi^2(5) = 10.131$, $p = .072$, CFI $= .992$, Gamma $= .995$, SRMR $= .018$.

Table 2
Goodness of fit indices for different models and range of factor loadings for measurement models.

Models	χ^2	*df*	*p*	CFI	Gamma	SRMR/WRMR[a]	BIC	λ_{range}
M01: MicroDYN (item)	66.379	53	.103	.996	1.000	.710[a]	–	[.26; .93]
M02: MicroDYN (parcel)	13.882	8	.085	.989	.991	.024	593	[.68; .83]
M03: MicroFIN (item)	78.260	53	.014	.886	.997	.798[a]	–	[.27; .77]
M04: MicroFIN: alternative (item)	13.473	9	.142	.956	.994	.036	2963	[.32; .56]
M05: CPS (parcel)	54.797	25	.001	.961	.976	.036	1715	[.55; .82]
M06: BIS (parcel)	80.351	71	.210	.993	.992	.042	4645	[.64; .85]
M07: figural reasoning (item)	3.972	5	.553	1.000	1.000	.023	3300	[.39; .75]
M08: BOWIT: two dimensions (item)	1048.486	944	.010	.834	.976	1.008[a]	–	[.03; .73]
M09: BOWIT: Unidimensional (item)	1126.222	945	.000	.711	.943	1.082[a]	–	[.03; .62]
H1.1: Intelligence → CPS (fom)	358.523	282	.001	.969	.976	.052	8880	
H1.2: Figural reasoning → CPS (fom)	105.005	74	.010	.969	.981	.039	4968	
H1.3: Intelligence + CPS (nfm)	394.883	282	.000	.954	.967	.066	8917	
H1.4: Intelligence + CPS (nfm; no specific CPS factor)	417.457	283	.000	.945	.958	.074	8934	
H1.5: Figural reasoning + CPS (nfm)	104.971	73	.008	.968	.980	.039	4973	
H1.6: Figural reasoning + CPS (nfm, no specific CPS factor)	124.795	74	.000	.949	.969	.062	4987	
H2.1: Intelligence → GPA (fom)	126.843	121	.340	.996	.997	.042	7641	
H2.2: Intelligence + CPS → GPA (fom)	381.248	302	.001	.969	.977	.052	9305	
H2.3: Figural reasoning → GPA (fom)	10.030	9	.348	.995	.997	.034	3748	
H2.4: Figural reasoning + CPS → GPA (fom)	129.346	86	.002	.958	.976	.043	5411	
H2.5: Intelligence + CPS → GPA (nfm)	416.834	302	.000	.954	.966	.065	9341	
H2.6: Figural reasoning + CPS → GPA (nfm)	127.933	85	.001	.957	.975	.043	5416	

Note. df = degrees of freedom; CFI = Comparative Fit Index; Gamma = Gamma Hat; SRMR = Standardized Root Mean Square Residual; WRMR = Weighted Root Mean Square Residual; BIC = Bayesian Information Criterion (only for MLR estimator); λ_{range} = range of factor loadings from observed indicators to latent factors; χ^2 and *df* estimates are based on MLR and WLSMV, respectively. M = Measurement models; H = Models according to our hypotheses, fom = first-order model, nfm = nested factor model; [a] = WRMR.

especially because developing a new measurement model requires a more comprehensive approach with several data sources (Schermelleh-Engel et al., 2003). Therefore, we calculated three parcels across the 45 items to reduce any noisy variance that was not assumed to be part of the representation of the latent construct (Little et al., 2013). A one-dimensional, just-identified measurement model showed significant and substantial factor loadings (all λs > .68).

3.2. Research issue 1: relation between CPS and established constructs of intelligence

Hypothesis 1.1. CPS and intelligence in the first-order factor model approach.

In our first analysis, we expected that less than 60% of the variance in CPS would remain unexplained when CPS was regressed on established intelligence constructs using the first-order factor model approach. The corresponding first model in which CPS was statistically predicted by a broad operationalization of intelligence showed a good fit (see Table 2, H1.1). Reasoning ($\beta = .85$, SE = .36, $p < .01$) and creativity ($\beta = -.34$, SE = .26, $p = .04$) were significant predictors of CPS, whereas mental speed ($\beta = -.02$, SE = .26, $p = .94$), memory ($\beta = .19$, SE = .21, $p = .16$), and general knowledge ($\beta = -.06$, SE = .17, $p = .56$) were not. The negative relation between creativity and CPS combined with the low zero-order correlation (see Table 3) indicated that creativity had a suppressor effect in this model (Pedhazur, 1997). Overall, 60.2% (95% CI [52.4, 68.0]) of the variance in CPS was explained, leaving 39.8% unexplained.

In the next model, we used a narrow operationalization of intelligence comprised of only figural reasoning to predict CPS so that we could compare our results with the previous findings from the distinctness perspective. The model fit was also good (see Table 2, H1.2). Figural reasoning ($\beta = .70$, SE = .21, $p < .01$) explained 48.6% (95% CI [39.4, 57.8]) of the variance in CPS and, thus, 51.4% of the variance remained unexplained.

In line with Hypothesis 1.1, less than 60% of the variance in CPS remained unexplained by established intelligence constructs. In fact, combining the advantages of previous studies led to only 39.8% of the variance in CPS remaining unexplained by established intelligence constructs, whereas a narrow operationalization of intelligence yielded a proportion of unexplained variance that was similar to the amount found in previous studies. In summary, Hypothesis 1.1 was supported.

Hypothesis 1.2. CPS and established intelligence constructs in the nested factor model approach.

Hypothesis 1.2 predicted that a specific CPS factor would emerge if CPS and established intelligence constructs were modeled in a nested factor model. We applied a nested factor model according to Eid, Lischetzke, Nussbeck, and Trierweiler's (2003) modifications. More specifically, rather than modeling specific factors for each ability, one specific factor was set as a reference construct and was thus not modeled as a specific factor. Consequently, the *g*-factor was explicitly determined by the variance shared between the reference construct and the specific abilities (Eid et al., 2003). In line with the prominent role of reasoning in intelligence research (e.g., Carroll, 1993, Wilhelm, 2005), it seemed adequate to define reasoning as the reference construct. In doing so, the *g*-factor was determined by reasoning, and the specific factors contained the unique systematic variance of each specific construct independent of reasoning.[7]

The resulting model (see Fig. 3) showed a good fit (see Table 2, H1.3). Factor loadings on the *g*-factor were all significant and substantial (*Mdn* $\lambda = .51$) except the one for creativity ($\lambda = .26$, $p < .01$). Factor loadings on the specific factors (including a distinct CPS factor) were all significant and substantial in the following order: general knowledge (*Mdn* $\lambda = .68$), memory (*Mdn* $\lambda = .62$), creativity (*Mdn* $\lambda = .59$), CPS (*Mdn* $\lambda = .59$), and mental speed (*Mdn* $\lambda = .53$). In addition, we calculated the specific reliabilities ω_s of the specific factors. The reliability was acceptable for the specific factors of memory ($\omega_s = .53$), creativity ($\omega_s = .56$), and general knowledge ($\omega_s = .58$), but it was below the recommended benchmark for mental speed ($\omega_s = .41$) and CPS ($\omega_s = .41$). To examine whether the existence of a CPS factor was justified, we compared this model with a model without a specific CPS factor (i.e., the CPS indicators loaded on the *g*-factor alone). This latter model showed an acceptable (see Table 2, H1.4) but significantly worse

[7] Please note that a nested factor model without Eid et al.'s (2003) modification provided generally similar findings. However, the specific factor for reasoning tended to collapse in some models (i.e., the factor loadings for the specific factor of reasoning decreased substantially).

Table 3
Descriptive statistics, reliability estimates, and correlations.

Measure	Min	Max	*M*	*SD*	ω	1)	2)	3)	4)	5)	6)	7)
1) Reasoning	42	120	79.42	16.01	.86	–	.70	.56	.56	.54	.72	.44
2) Mental speed	101	338	241.86	37.62	.80	.57	–	.66	.64	.29	.47	.17
3) Memory	54	147	99.16	17.84	.77	.48	.53	–	.43	.31	.49	.37
4) Creativity	56	161	94.44	18.77	.79	.47	.52	.34	–	.18	.20	.30
5) General knowledge	7	41	20.74	6.18	.77	.46	.23	.26	.15	–	.39	.23
6) CPS	−3.1	1.54	0.00	0.85	.87	.51	.32	.35	*.10*	.29	–	.31
7) GPA	3	6	5.20	0.64	–	.41	*.13*	.33	.25	.20	.24	–

Note. Descriptive statistics are based on single items and their total scores. CPS is based on z-values. GPA was reverse-coded as 6 = excellent, 1 = insufficient. ω = McDonald's Omega (internal consistency). Manifest Pearson correlations (based on total scores) are reported below and latent correlations above the principal diagonal. Nonsignificant correlations are written in italics.

model fit: $\Delta\chi^2 = 27.866$, $df = 1$, $p < .01$; ΔBIC = 17. Therefore, a model with a specific CPS factor fit the data better.

To ensure that we could explicitly compare our findings with previous studies from the distinctness perspective, the second nested factor model involved figural reasoning (as a narrow operationalization of intelligence) and CPS. The model showed a good fit (see Table 2, H1.5). Figural reasoning was set as the reference construct, and thus, the model contained a *g*-factor and only one specific CPS factor. Factor loadings were all significant and substantial (*Mdn* λ = .58 for the *g*-factor; *Mdn* λ = .59 for the specific CPS factor). The specific reliability ω_s of the specific CPS factor was again below the recommended benchmark ($\omega_s = .41$). For a model without a specific CPS factor (which is equivalent to a one-dimensional model with a single latent factor for figural reasoning and CPS), the model fit was acceptable (see Table 2, H1.6) but significantly worse: $\Delta\chi^2 = 26.201$, $df = 1$, $p < .01$; ΔBIC = 14. Thus, a model with a specific CPS factor fit the data better.

In summary, the findings supported Hypothesis 1.2. A *g*-factor based on reasoning could not fully explain the variance in CPS. In fact, a specific CPS factor—independent of the *g*-factor and other specific factors—accounted for unique variance in the CPS measures. However, the specific reliability of the CPS factor was rather weak. The findings were independent of the breadth of the operationalization of intelligence. Hence, Hypothesis 1.2 was supported.

3.3. Research issue 2: incremental predictive validity of CPS

Hypothesis 2.1. CPS and GPA in the first-order factor model approach.

Hypothesis 2.1 predicted that CPS would be positively related to GPA and would incrementally explain variance in GPA beyond established intelligence constructs when the first-order factor model approach was used. In order to examine this hypothesis, we used a baseline model in which only established constructs of intelligence predicted GPA. In a subsequent model, we used CPS as an additional predictor of GPA to examine the incremental predictive validity of CPS beyond intelligence constructs.

When a broad operationalization of intelligence was used, the baseline model showed a good fit (see Table 2, H2.1). GPA was statistically significantly predicted by reasoning (β = .52, SE = .09, $p < .01$), memory (β = .37, SE = .07, $p < .01$), creativity (β = .22, SE = .06, $p = .01$), and mental speed (β = −.58, SE = .10, $p < .01$), but not by general knowledge (β = −.05, SE = .07, $p = .33$). The negative relation between mental speed and GPA in combination with the low zero-order correlation between them (see Table 3) indicated that mental speed functioned as a suppressor in this model (Pedhazur, 1997). The amount of variance explained in school grades was 31.9% (95% CI [22.2, 41.6]). In the subsequent model with a good fit (see Table 2, H2.2), established intelligence constructs and CPS simultaneously predicted GPA (see Fig. 4). As a result, only reasoning (β = .57, SE = .15, $p < .01$), mental speed (β = −.58, SE = .10, $p < .01$; again as a suppressor), and memory (β = .39, SE = .07, $p < .01$) were significant predictors, but creativity (β = .20, SE = .08, $p = .05$), general knowledge (β = −.05, SE = .07, $p = .32$), and CPS (β = −.04, SE = .08, $p = .41$) were not. The amount of variance explained in GPA was 32.6% (95% CI [22.9, 42.3]), indicating no incremental predictive power of CPS beyond established intelligence constructs: $\Delta R^2 = 0.7\%$; $F(1, 220) = 2.28$, $p = .13$.

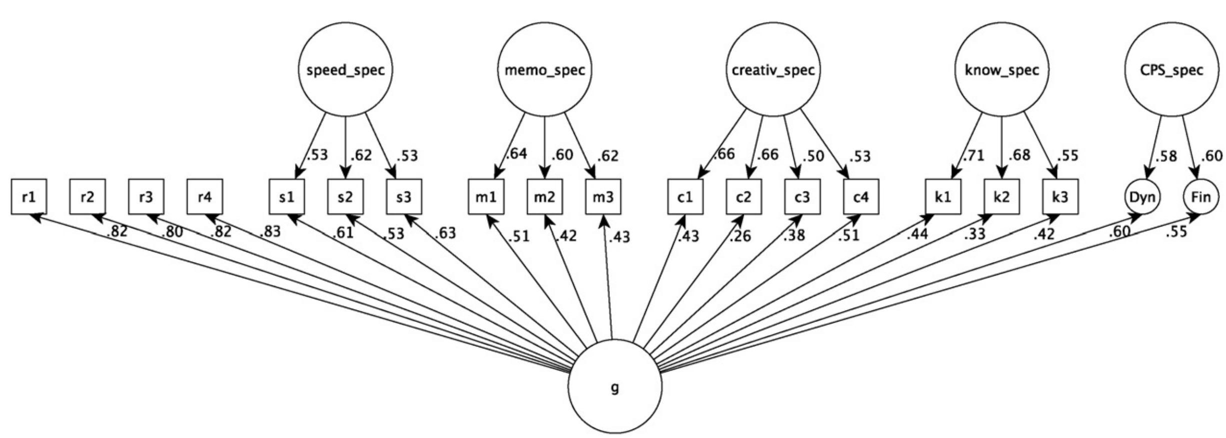

Fig. 3. Nested factor model for established intelligence constructs and CPS based on Eid et al.'s (2003) modifications. Reasoning was set as the reference construct. *g*: general factor of intelligence (*g*-factor); speed_spec: specific mental speed ability; memo_spec: specific memory ability; creativ_spec: specific creativity ability; know_spec: specific general knowledge ability; CPS_spec: specific CPS ability; r1–r4: parcel scores for reasoning items; s1–s4: parcel scores for mental speed items; m1–m3: parcel scores for memory items; c1–c4: parcel scores for creativity items; k1–k3: parcel scores for general knowledge items; Dyn: latent factor for MicroDYN; Fin: latent factor for MicroFIN. Standardized model solution is shown. Error terms are not depicted.

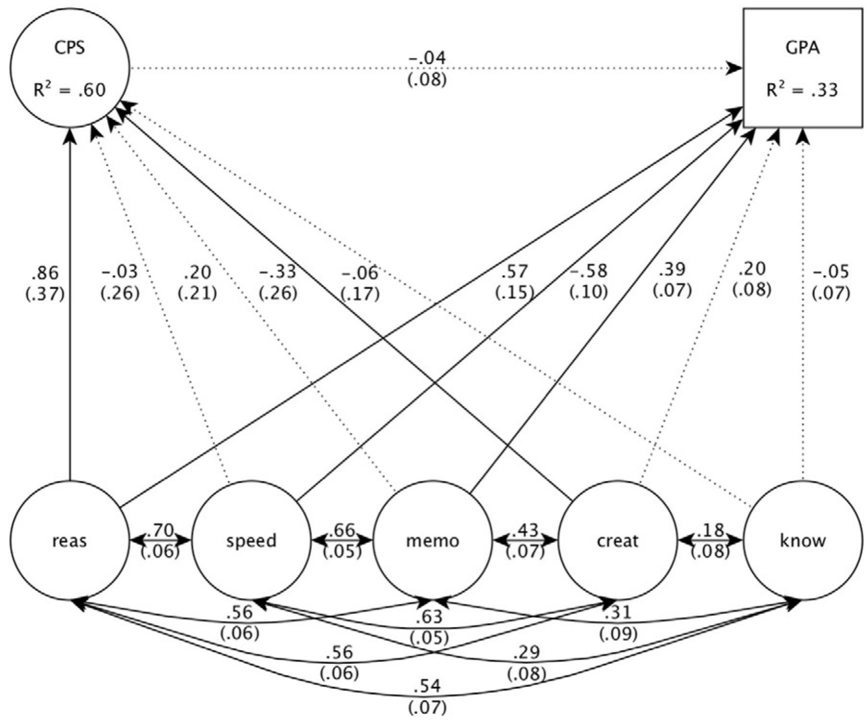

Fig. 4. Structural model including established intelligence constructs, CPS, and GPA in the first-order factor model approach. Reas: reasoning; speed: mental speed; memo: memory; creativ: creativity; know: general knowledge. The standardized model solution is shown with standard errors in parentheses. Dashed lines indicate nonsignificant relations. Manifest variables are not displayed.

To compare these findings with the ones obtained in previous studies from the distinctness perspective, these analyses were repeated with a narrow operationalization of intelligence. In the baseline model, only figural reasoning predicted GPA. The model showed a good fit (see Table 2, H2.3), and figural reasoning was a statistically significant predictor of GPA (β = .25, SE = .04, $p < .01$), explaining 6.2% (95% CI [0.2, 12.2]) of the variance. In the subsequent model, figural reasoning and CPS predicted GPA simultaneously. The model also showed a good fit (see Table 2, H2.4). It is interesting that only CPS (β = .25, SE = .07, $p = .04$) was significantly positively related to GPA, but figural reasoning was not (β = .09, SE = .08, $p = .24$). However, figural reasoning and CPS explained 10.1% (95% CI [2.8, 17.4]) of the variance in GPA; that is, $\Delta R^2 = 3.9\%$ of the variance in GPA was incrementally predicted by CPS beyond figural reasoning. This difference in explained variance was statistically significant, $F(1224) = 9.78$, $p < .01$.

In general, the findings only partly supported the prediction that CPS would incrementally explain variance in school grades beyond established constructs of intelligence. Using the first-order factor model approach, CPS incrementally explained variance in GPA beyond established intelligence constructs only when figural reasoning was used. However, no significant incremental variance was explained when a broad operationalization of intelligence was considered. Therefore, Hypothesis 2.1 was not supported.

Hypothesis 2.2. CPS and GPA in the nested factor model approach.

Hypothesis 2.2 predicted that CPS would incrementally explain variance in school grades beyond established intelligence constructs in a nested factor model approach. Therefore, we applied the nested factor model from Hypothesis 1.2 to predict GPA. The model utilizing a broad operationalization of intelligence showed a good fit (see Table 2, H2.5). GPA was significantly predicted by the *g*-factor (β = .43, SE = .04, $p < .01$) and the specific factors of mental speed (β = −.31, SE = .05, $p < .01$), memory (β = .21, SE = .05, $p < .01$), and creativity (β = .14, SE = .05, $p = .03$) but not by the specific factors of general knowledge (β = −.02, SE = .06, $p = .42$) and CPS (β = −.02, SE = .07, $p = .43$). This model explained 34.4% (95% CI [24.7, 44.1]) of the variance in GPA.

Similar to the results for Hypothesis 2.1, the findings changed when only figural reasoning was used to operationalize intelligence. The corresponding nested factor model showed a good fit (see Table 2, H2.6). In this model, the *g*-factor (β = .26, SE = .04, $p < .01$) and the specific CPS factor (β = .18, SE = .07, $p = .04$) were both positively and significantly related to GPA. Figural reasoning and CPS explained a total of 10.1% (95% CI [2.8, 17.4]) of the variance in GPA.

In summary, CPS incrementally predicted school grades beyond established intelligence constructs only when intelligence was operationalized narrowly. A broad operationalization of intelligence led to an absence of an incremental prediction of school grades by the specific CPS factor. Thus, Hypothesis 2.2 was not supported.

4. Discussion

The present study examined the role of CPS in the nomological network of intelligence and, thus, whether CPS should be seen as a distinct cognitive construct or not. For the first time in CPS research, we combined the advantages of several studies from two different perspectives (i.e., redundancy perspective vs. distinctness perspective). In detail, we used two psychometrically sound CPS assessment tools and compared a broad and a narrow operationalization of intelligence to investigate the relation between CPS and established intelligence constructs. Moreover, we examined the incremental predictive validity of CPS beyond established intelligence constructs with regard to an external criterion (i.e., school grades). For all analyses, we applied different measurement models (i.e., first-order factor model vs. nested factor model) to ensure the robustness of our findings.

4.1. Relations between CPS and established constructs of intelligence

In line with previous research, we found a strong relation between CPS and established intelligence constructs. More precisely, only 39.8%

of the variance in CPS could not be explained by established intelligence constructs in the first-order factor model approach. The proportion of unexplained variance was substantially smaller in comparison with earlier findings in which up to 58% of the variance was unexplained when a broad operationalization of intelligence in combination with suboptimal psychometric CPS assessment tools were used from the redundancy perspective (e.g., Süß, 1996). It was also smaller than the value of up to 60% of the variance in CPS that was unexplained when a narrow operationalization of intelligence in combination with state-of-the-art CPS assessment tools were used in studies from the distinctness perspective (e.g., Wüstenberg et al., 2012). Therefore, the present findings demonstrate the impact and importance of our approach in which we combined the advantages of previous studies from both perspectives in order to obtain a more comprehensive view of the relation between CPS and established intelligence constructs. More specifically, our findings suggest that future studies should use a construct-valid operationalization of intelligence and psychometrically sound CPS assessment tools when their relation on a construct level is the primary research question.

It is noteworthy that the findings from the first-order factor model approach provided only limited information with regard to whether CPS should be considered a distinct construct or not. In fact, a relation similar in magnitude to the one found in our study (i.e., 51.4% of the variance left unexplained with a narrow operationalization of intelligence) was interpreted to indicate either convergent (e.g., Kröner et al., 2005, Süß, 1996) or discriminant (e.g., Wüstenberg et al., 2012) validity in different studies. Even our finding of 39.8% of the variance in CPS left unexplained (with the broad operationalization of intelligence) could be interpreted as substantially different from zero (i.e., equivalent to the often used criterion $r = 1.00$ for construct identity; e.g., Neubert et al., 2015a; Sonnleitner et al., 2013). Therefore, this finding could be taken as evidence for a distinct construct or as sufficient to assume construct identity because different intelligence tests usually do not share more common variance (e.g., Heller, Kratzmeier, & Lengfelder, 1998). Thus, it seems to be difficult to define an unambiguous criterion for determining whether two latent factors represent the same construct in the first-order factor model approach[8]—although the vast majority of previous CPS research has been based on that approach.

Overcoming the shortcomings of the first-order factor model approach, the nested factor model approach seems to be more appropriate in this context. In this approach, specific factors representing "pure" variance independent of a *g*-factor and other specific factors can be examined and evaluated according to objective criteria (Murray & Johnson, 2013). Our analyses with the nested factor model approach showed a specific CPS factor in addition to the established intelligence construct factors of reasoning, mental speed, memory, creativity, and general knowledge (i.e., crystallized intelligence). Therefore, our results supported the notion that CPS measures cover specific, systematic variance that is independent of established intelligence constructs. Further analyses of the reliability of the specific CPS factor (i.e., omega specific) demonstrated that 41% of the total variance in CPS measures was independent of the *g*-factor. This general pattern of findings was independent of whether a broad or narrow operationalization of intelligence was used. Although a magnitude of at least 50% was recommended for a meaningful interpretation of construct-related variance (Reise et al., 2013), similar findings for the established constructs in the BIS model (i.e., from 41% to 56%; see also Brunner & Süß, 2005) emphasize the preliminary nature of that recommendation.

In summary, our findings on the relations of established intelligence constructs and CPS as highlighted in Research Issue 1 support the distinctness perspective by indicating that the construct of CPS can be separated from established intelligence constructs. Even though the amount of variance in CPS left unexplained was lower than in previous studies, and its specific factor demonstrated rather low reliability, it seems that measures of CPS cover a specific ability that is not well-incorporated in traditional tests of intelligence (Greiff, Fischer, et al., 2013; Sonnleitner et al., 2013; Wüstenberg et al., 2012).

Although not the focus of this study, the cognitive requirements for CPS were further illuminated by our results for Hypothesis 1.1 (see Fig. 4). Whereas reasoning played a major role in differentiating successful from unsuccessful complex problem solvers, mental speed, memory, creativity, and general knowledge did not help to differentiate between problem solvers of different proficiency (based on the operationalization used in the present study). This pattern is in line with previous findings (based on traditional CPS assessment tools; see Süß, 1996; Wittmann & Süß, 1999) in which reasoning showed the highest relation to CPS performance compared with other intelligence constructs.

However, our results provide only preliminary insights into the cognitive requirements for CPS. More fine-grained empirical research on the cognitive processes of CPS—in particular, in distinguishing it from established intelligence constructs—is rare (for a theoretical discussion, see, e.g., Fischer et al., 2012, or Greiff, Fischer, et al., 2014). Therefore, future studies should examine the cognitive requirements that make CPS a distinct construct in the nomological network of intelligence. In this respect, future research could, for example, investigate the relations between CPS and cognitive learning processes (see Guthke, 1982), which are considered highly relevant for solving complex problems (Greiff, Fischer, et al., 2014). There are only a few studies with the above-mentioned shortcomings that have reported modest (Beckmann, 1994) or no substantial (Süß, 2001) correlations between intelligence tests that cover learning processes and CPS. Therefore, a replication of the present findings with psychometrically sound CPS assessment tools and a broad operationalization of intelligence that includes the assessment of learning processes could provide further evidence for the distinctness perspective.

4.2. Incremental predictive validity of CPS beyond established intelligence constructs

The second important issue in investigating the construct validity of CPS was the incremental predictive validity of CPS. If CPS is a distinct construct, then it should incrementally explain external criteria (e.g., school grades) beyond established intelligence constructs. In line with a variety of previous studies (e.g., Greiff, Wüstenberg, et al., 2013; Kretzschmar et al., 2014) that support the distinctness perspective, we demonstrated the incremental predictive validity of CPS beyond a narrow operationalization of intelligence in explaining variance in school grades. However, no incremental validity of CPS was found if a broad operationalization of intelligence was used. Furthermore, the general pattern of findings was independent of the applied measurement model approach (i.e., first-order factor model vs. nested factor model). Therefore, the breadth of the operationalization of intelligence seems to have an important impact on the incremental predictive validity of CPS beyond established intelligence constructs.

A possible explanation for the impact of the breadth of the operationalization of intelligence is the notion of the Brunswik symmetry principle (Wittmann & Süß, 1999). The basic idea underlying Brunswik symmetry is that correlations are reduced if unequal levels of aggregation (e.g., in terms of a hierarchy of cognitions) are used. For example, GPA covers a variety of different cognitive (and noncognitive) processes in several domains (e.g., math, languages, arts) and, thus, it can be considered rather highly aggregated. On the other hand, figural reasoning is a very specific and, hence, not-so-aggregated construct in the nomological network of intelligence (Wittmann & Hattrup, 2004). Because of these unequal levels of aggregation between

[8] Please note that the multitrait multimethod approach (Campbell & Fiske, 1959; for an application in CPS reserach, see Greiff, Fischer, et al., 2013) also does not provide such a criterion other than to compare the relations between convergent and discriminant measures. The decision of whether a measure is used for convergent or discriminant validity depends on the researcher.

predictor (i.e., figural reasoning) and criterion (i.e., GPA), the relation between them is reduced, and an additional cognitive predictor can show additional predictive power (e.g., CPS). In contrast to figural reasoning, a broad operationalization of intelligence can be assumed to be highly aggregated (Wittmann & Hattrup, 2004). Thus, predictor and criterion are more symmetrical, and the correlation is not reduced by an unequal aggregation level. As a consequence, there is less variance in the criterion left unexplained, leading to an absence of incremental validity offered by CPS beyond established intelligence constructs. In conclusion, it seems promising to consider Brunswik symmetry when examining and interpreting the relations between CPS and other constructs as well as external criteria.

On the basis of our findings, we can conclude that CPS did not show incremental validity in predicting school grades beyond established intelligence constructs. This means that the distinctness perspective was not supported with regard to incremental predictive validity. In fact, previous findings on the incremental prediction of school grades beyond narrow operationalizations of intelligence (i.e., figural reasoning) might be the result of methodological effects with respect to the Brunswik symmetry principle.

4.3. Limitations

There are several limitations to our study, and we would like to emphasize three crucial issues. To our knowledge, this is the first study to examine the relation between general knowledge (i.e., crystallized intelligence) and CPS operationalized with state-of-the art CPS measurement tools (i.e., multiple-task tests). Our findings indicate that general knowledge plays only a minor role in current operationalizations of CPS (cf. Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2015). Although general knowledge was assessed with a common approach (i.e., verbal knowledge tasks), previous research has shown that the domination of verbal tasks in the measurement of crystallized intelligence might lead to a confounding of knowledge and verbal abilities (Beauducel, Brocke, & Liepmann, 2001). In addition, the psychometric quality of the applied measure (BOWIT) should be examined further. Therefore, conclusions with regard to general knowledge or crystallized intelligence, respectively, should be drawn only with caution and replicated with more sophisticated measures.

With regard to external criteria, the sufficiency of GPA as a criterion for CPS could be questioned. As noted above, GPA is a conglomerate of a variety of processes in different domains. And although CPS is generally considered to be domain-general, there is some evidence that CPS is more relevant in the science domain than in languages and art (e.g., Greiff et al., 2015, Kretzschmar et al., 2014). Hence, GPA and CPS might not be sufficiently symmetrical in terms of the Brunswik symmetry principle, and thus, the correlation between them might be reduced. Future studies should also consider the potential differential importance of CPS in different domains (e.g., Greiff, Fischer, et al., 2013). Furthermore, an examination of the incremental validity of CPS should rely on more than just school grades as a criterion. For example, CPS is considered important in the occupational context (Neubert, Mainert, Kretzschmar, & Greiff, 2015b), making professional success a promising external criterion for CPS (see e.g., Danner, Hagemann, Holt, et al., 2011).

Finally, as in previous CPS studies (e.g., Greiff, Fischer, et al., 2013; Greiff & Fischer, 2013; Süß, 1999; Wüstenberg et al., 2012; but for more heterogeneous samples, see e.g., Greiff, Wüstenberg, et al., 2013; Sonnleitner et al., 2013; Süß, 1996), a sample of university students was used. For most German majors, admission depends on GPA, and as a result, our sample consisted of students with above-average cognitive and academic performances. Although our sample (GPA: $M = 1.8$, $SD = 0.64$; original scale) is comparable to samples used in previous CPS studies (e.g., GPA: $M = 1.7$, $SD = 0.7$; Wüstenberg et al., 2012), such homogenous samples limit the generalizability of our findings. In fact, homogenous samples are characterized by lower variance and, thus, tend to provide reduced correlations (Kline, 1994). Consequently, the number of factors in factor analyses might be arbitrarily increased, and the relations between CPS, established intelligence constructs, and GPA might be arbitrarily decreased. Thus, we highly encourage other researchers to replicate our findings with more heterogeneous samples.

5. Conclusions

Since the beginning of CPS research (e.g., Dörner et al., 1983), researchers have questioned whether CPS is different from established intelligence constructs and, thus, whether current theories of intelligence should be extended by adding a CPS construct or even whether the notion of a construct independent from intelligence is warranted. The present study aimed to shed light on this controversial issue by combining several advantages of previous studies and, thus, to present a profound view of CPS in the nomological network of intelligence. Our findings provide evidence for both the redundancy and distinctness perspectives of CPS. First, we found unique and systematic variance in CPS independent of established intelligence constructs. However, we did not find that CPS offered incremental validity in predicting school grades beyond established intelligence constructs.

Of course, future studies will have to replicate our findings to provide further support for these ideas. In this respect, the present study offers important points that researchers would be wise to consider when investigating the construct validity of CPS in the future. First, a narrow operationalization of intelligence (e.g., figural reasoning) is not sufficient for examining the role of CPS in the broader nomological network of intelligence. Instead, a comprehensive operationalization (i.e., at least a broad operationalization of reasoning) is necessary for investigating the construct validity of CPS. Second, as measurement models including a *g*-factor (e.g., the nested factor measurement model) better reflect contemporary perspectives on the structure of intelligence (Brunner, 2008), studies targeting CPS, established intelligence constructs, and their unique contributions should include such attempts at modeling. What is more, the nested factor model provides more clearly separable criteria for evaluating a distinct CPS construct compared with first-order factor models. Therefore, we highly recommend the use of nested factor models in future research that aims to address the question of whether CPS is a distinct construct. Finally, neither the correlation between established intelligence constructs and CPS nor an examination of the incremental predictive validity alone is sufficient for arguing for or against a distinct CPS construct. The combination of both, optimally combined with diverse external criteria (e.g., from professional settings, Neubert et al., 2015b), should be considered. Considering these issues in future studies would provide strong evidence for answering the question of whether CPS should be included as a distinct construct in contemporary theories of intelligence. Most important of all, we think one of the next crucial steps in CPS research should be to face the scientific challenge of theoretically explaining and empirically verifying the unique cognitive aspects of CPS in comprehensive theories of intelligence.

Appendix B. Detailed descriptions of the utilized MicroFIN tasks

B.1. Task 1: Concert-o-maton

The “Concert-o-maton” has three input variables. They are labeled “music group” (two qualitatively different values) and “stage” (two values). In addition, the factor “admission fee” can be varied (three ordinal values). The outcome variable is the number of visitors (three ordinal values). All three input variables can be varied independently. Based on the combination of “music group” and “stage,” the influence of “admission fee” on the number of visitors is varied according to specific rules (e.g., reversed effect of admission fee for one combination; no influence of admission fee on the output variable at all in other combinations).

Appendix A

Table A1
MicroDYN tasks characteristics: Linear structural equations, system size, and type of effects for the MicroDYN tasks.

Task	Linear structural equations	System size	Effects
1: Wind engine	$X_{t+1} = 1 * X_t + 2 * A_t + 2 * B_t + 0 * Ct$ $Y_{t+1} = 1 * Y_t + 0 * A_t + 0 * B_t + 2 * Ct$	2 × 3 - System	Only effects of inputs
2: Factory	$X_{t+1} = 1 * X_t + 2 * A_t + 2 * B_t + 0 * C_t$ $Y_{t+1} = 1 * Y_t + 0 * A_t + 2 * B_t + 2 * C_t$ $Z_{t+1} = 1 * Z_t + 0 * A_t + 0 * B_t + 2 * C_t$	2 × 3 - System	Only effects of inputs
3: Logistic	$X_{t+1} = 1 * X_t + 2 * A_t + 0 * B_t + 0 * C_t$ $Y_{t+1} = 1 * Y_t + 0 * A_t + 2 * B_t + 0 * C_t + 3$	2 × 3 - System	Effects of inputs and outputs
4: Research	$X_{t+1} = 1 * X_t + 0 * A_t + 0 * B_t + 0 * C_t$ $Y_{t+1} = 1 * Y_t + 2 * A_t + 2 * B_t + 0 * C_t + 3$ $Z_{t+1} = 1 * Z_t + 0 * A_t + 0 * B_t + 2 * C_t$	3 × 3 - System	Effects of inputs and outputs
5: Team leading	$X_{t+1} = 1 * X_t + 2 * A_t + 0 * B_t + 0 * C_t$ $Y_{t+1} = 1 * Y_t + 2 * A_t + 2 * B_t + 0 * C_t + 3$ $Z_{t+1} = 1 * Z_t + 0 * A_t + 0 * B_t + 2 * C_t$	3 × 3 - System	Effects of inputs and outputs
6: Employee selection	$X_{t+1} = 1.33 * X_t + 2 * A_t + 2 * B_t + 0 * C_t$ $Y_{t+1} = 1 * Y_t + 0 * A_t + 2 * B_t + 0 * C_t$ $Z_{t+1} = 1 * Z_t + 0 * A_t + 2 * B_t + 2 * C_t$	3 × 3 - System	Effects of inputs and outputs

Note. Linear structural equations: values of the output variables (X, Y, Z) at time point t + 1 depending on input (A, B, C) and output variables at time point t. System size: number of input and output variables. Effects: only effects of input variables on output variables or effects of input and output variables on output variables (i.e., linear or exponential Eigendynamics).

B.2. Task 2: Exchange-o-maton

The “Exchange-o-maton” simulates the exchange of gifts between two persons, their well-being (three ordinal values per person) depending on the gifts they receive (four gifts), as well as the well-being of the interaction partner. Each person's well-being is increased by only one type of gift and decreased by another type. Additional rules have to be explored by the problem solver: The highest level of well-being can be achieved only if the other person's well-being is at least at a certain level. If the well-being of one person drops to the lowest level, the other person's well-being is reduced to a medium level if it was on the highest level before.

B.3. Task 3: Fish-o-maton

The “Fish-o-maton” features three ordinal input variables, each with four input values. The output variable, an aquarium, has five possible values: Empty, soiled, a few fish, a moderate number of fish, and many fish (nominal variable with ordinal elements). Different values in the output variable are shown to depend on the input variables and their relations to each other (e.g., when all input variables are set to equal levels, fish appear in the aquarium, the number depending on the level of the input values).

B.4. Task 4: Plan-o-maton

In the “Plan-o-maton”, four different buildings (i.e., input values) are presented in a 2 × 2 matrix format. The position of the values can be exchanged by pressing one of four buttons situated next to the matrix (one for each row/column, e.g., the one on the side of the bottom row changing the position of the picture from the bottom left to the bottom right, and vice versa). There are four output variables, each with two values that are presented between the pictures. The output values are based on the combination of the respective values in the cells of the matrix (i.e., pairs of the buildings are matched). The combination of two pictures leads to one of two output values independent of where the pictures are shown.

B.5. Task 5: Green-o-maton

The “Green-o-maton” requires participants to explore and control a greenhouse. The greenhouse produces pumpkins of different sizes and colors (one output variable with four values), depending on the settings of three input variables, which can be manipulated independently. One of the input variables controls the temperature (four values representing two temperature levels), and the other two determine the use of two fertilizers (two binary input variables). Pumpkin color is determined by temperature level alone and size is a result of the combination of fertilizers for a given temperature level (i.e., only one fertilizer produces large pumpkins at each temperature level).

B.6. Task 6: Cook-o-maton

The “Cook-o-maton” simulates the creation of a soup (one output variable) with the help of three abstract ingredients (i.e., three input variables with two values each). To operate the automaton, the problem solver has to select from the ingredients (binary selection of each ingredient) and press a start button. Depending on the selected ingredients and specific combination rules (e.g., the omission of water leading to burnt ingredients), one of three soups or burnt ingredients is displayed in a pot (four different output values after the first run). In addition, adding the third ingredient in a second step and pressing the start button again further modifies one of the resulting output values (additional output value in the second run), so changes in the problem environment have to be explored as well as combination rules.

References

Amthauer, R., Brocke, B., Liepmann, D., & Beauducel, A. (2001). *I-S-T 2000 R - Intelligenz-Struktur-Test 2000 R [Intelligence Structure Test 2000 (revised)].* Göttingen: Hogrefe.

Bahrick, H. P., Hall, L. K., & Berger, S. A. (1996). Accuracy and distortion in memory for high school grades. *Psychological Science*, *7*(5), 265–271 http://doi.org/10.1111/j.1467-9280.1996.tb00372.x.

Beauducel, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment*, *18*(2), 97–112 http://doi.org/10.1027//1015-5759.18.2.97.

Beauducel, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: Facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences*, *30*(6), 977–994.

Beauducel, A., Liepmann, D., Felfe, J., & Nettelnstroth, W. (2007). The impact of different measurement models for fluid and crystallized intelligence on the correlation with personality traits. *European Journal of Psychological Assessment*, *23*(2), 71–78 http://doi.org/10.1027/1015-5759.23.2.71.

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

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 http://doi.org/10.1007/s11251-013-9280-7.

Brunner, M. (2008). No g in education? *Learning and Individual Differences*, *18*(2), 152–165 http://doi.org/10.1016/j.lindif.2007.08.005.

Brunner, M., & Süß, H. -M. (2005). Analyzing the reliability of multidimensional measures: An example from intelligence research. *Educational and Psychological Measurement*, *65*(2), 227–240 http://doi.org/10.1177/0013164404268669.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality*, *80*(4), 796–846 http://doi.org/10.1111/j.1467-6494.2011.00749.x.

Buchner, A., & Funke, J. (1993). Finite-state automata: Dynamic task environments in problem-solving research. *The Quarterly Journal of Experimental Psychology Section A*, *46*(1), 83–118 http://doi.org/10.1080/14640749308401068.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105http://doi.org/http://dx.doi.org/10.1037/h0046016

Carroll, J. B. (1993). *Human cognitive abilities. A survey of factoranalytic studies.* Cambridge: Cambridge University Press.

Core Team, R. (2013). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing Retrieved from http://www.R-project.org/.

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

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

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.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological Methods*, *8*(1), 38–60 http://doi.org/10.1037/1082-989X.8.1.38.

Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, *42*(3), 509–529 http://doi.org/10.1080/00273170701382864.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *The Journal of Problem Solving*, *4*(1), 19–41 http://doi.org/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 http://doi.org/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: Lawrence Erlbaum Associates.

Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle [Control of dynamic systems by building up and using subjective causal models]. *Zeitschrift für Psychologie*, *193*, 435–457.

Funke, J. (1992a). Dealing with dynamic systems: Research strategy, diagnostic approach and experimental results. *The German Journal of Psychology*, *16*(1), 24–43.

Funke, J. (1992b). *Wissen über dynamische Systeme: Erwerb, Repräsentation und Anwendung [Knowledge about dynamic systems: Acquisition, representation, and use].* Berlin; New York: Springer.

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 and Reasoning*, *7*(1), 69–89 http://doi.org/10.1080/13546780042000046.

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*, 133–142 http://doi.org/10.1007/s10339-009-0345-0.

Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence*, *52*, 71–79 http://doi.org/10.1016/j.intell.2015.07.006.

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

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, *24*(1), 13–23.

Greiff, S. (2012). *Individualdiagnostik der komplexen Problemlösefähigkeit [Assessment of problem solving ability].* Münster: Waxmann.

Greiff, S., & Fischer, A. (2013). Der Nutzen einer komplexen Problemlösekompetenz: Theoretische Überlegungen und empirische Befunde [The value of complex problem solving competency: Theoretical considerations and empirical results]. *Zeitschrift für Pädagogische Psychologie*, *27*(1), 27–39 http://doi.org/10.1024/1010-0652/a000086.

Greiff, S., Kretzschmar, A., Müller, J. C., Spinath, B., & Martin, R. (2014c). 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 http://doi.org/10.1037/a0035426.

Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2014a). Assessing complex problem-solving skills with multiple complex systems. *Thinking and Reasoning*, *21*(3), 356–382 http://doi.org/10.1080/13546783.2014.989263.

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013a). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence*, *41*(5), 579–596 http://doi.org/10.1016/j.intell.2013.07.012.

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 http://doi.org/10.1016/j.intell.2015.02.007.

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: a new assessment perspective. *Applied Psychological Measurement*, *36*(3), 189–213 http://doi.org/10.1177/0146621612439620.

Greiff, S., Wüstenberg, S., Molnar, G., Fischer, A., Funke, J., & Csapo, B. (2013b). Complex problem solving in educational contexts—something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, *105*(2), 364–379 http://doi.org/10.1037/a0031856.

Guthke, J. (1982). The learning test concept—an alternative to the traditional static intelligence test. *German Journal of Psychology*, *6*(4), 306–324.

Heller, K. A., Kratzmeier, H., & Lengfelder, A. (1998). *Matrizen-Test-Manual. Ein Handbuch zu den Standard Progressive Matrices von J.C. Raven [Matrices Test Manual. A Handbook of Standard Progressive Matrices of J.C. Raven].* Göttingen: Beltz-Testgesellschaft.

Hossiep, R., & Schulte, M. (2008). *BOWIT. Bochumer Wissenstest [BOWIT. Bochumer Knowledge Test].* Göttingen: Hogrefe.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55 http://doi.org/10.1080/10705519909540118.

Jäger, A. O. (1984). Intelligenzstrukturforschung: Konkurrierende Modelle, neue Entwicklungen, Perspektiven. [Structural research on intelligence: competing models, new developments, perspectives]. *Psychologische Rundschau*, *35*(1), 21–35.

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

Jensen, A. R., & Wang, L. -J. (1994). What is a good g? *Intelligence*, *18*, 231–258.

Kersting, M. (2001). Zur Konstrukt- und Kriteriumsvalidität von Problemlöseszenarien anhand der Vorhersage von Vorgesetztenurteilen über die berufliche Bewährung [On the construct and criterion validity of problem-solving scenarios based on the prediction of supervisor assessment of job performance]. *Diagnostica*, *47*(2), 67–76 http://doi.org/10.1026//0012-1924.47.2.67.

Kline, P. (1994). *An easy guide to factor analysis.* London: Routledge.

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 http://doi.org/10.1024/1010-0652/a000137.

Kröner, S. (2001). *Intelligenzdiagnostik per Computersimulation [Intelligence assessment via computer simulation].* Münster: Waxmann.

Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, *33*(4), 347–368 http://doi.org/10.1016/j.intell.2005.03.002.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*(1), 148–161 http://doi.org/10.1037/0022-3514.86.1.148.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*(2), 151–173 http://doi.org/10.1207/S15328007SEM0902_1.

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*(3), 285–300 http://doi.org/10.1037/a0033266.

Lu, L., Weber, H. S., Spinath, F. M., & Shi, J. (2011). Predicting school achievement from cognitive and non-cognitive variables in a Chinese sample of elementary school children. *Intelligence*, *39*(2–3), 130–140 http://doi.org/10.1016/j.intell.2011.02.002.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10 http://doi.org/10.1016/j.intell.2008.08.004.

Milling, P. M. (1996). Modeling innovation processes for decision support and management simulation. *System Dynamics Review*, *12*(3), 211–234.

Moshagen, M., & Musch, J. (2014). Sample size requirements of the robust weighted least squares estimator. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *10*(2), 60–70 http://doi.org/10.1027/1614-2241/a000068.

Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, *41*(5), 407–422 http://doi.org/10.1016/j.intell.2013.06.004.

Nachtigall, C., Kroehne, U., Funke, F., & Steyer, R. (2003). Pros and cons of structural equation modeling. *Methods of Psychological Research Online*, *8*(2), 1–22.

Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2015a). Extending the assessment of complex problem solving to finite state automata: Embracing heterogeneity. *European Journal of Psychological Assessment*, *31*(3), 181–194 http://doi.org/10.1027/1015-5759/a000224.

Neubert, J. C., Mainert, J., Kretzschmar, A., & Greiff, S. (2015b). The assessment of 21st century skills in industrial and organizational psychology: complex and collaborative problem solving. *Industrial and Organizational Psychology*, *18*(2), 238–268. http://dx.doi.org/10.1017/iop.2015.14.

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

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. Fort Worth: Harcourt Brace College Publishers.

Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg [On the relationship between test intelligence and success in problem solving]. *Zeitschrift für Psychologie*, *189*(1), 79–100.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–164.

Raven, J. (1958). *Advanced progressive matrices* (2nd ed.). London: Lewis.

Reeve, C. L., & Blacksmith, N. (2009). Identifying g: A review of current factor analytic practices in the science of mental abilities. *Intelligence*, *37*(5), 487–494 http://doi.org/10.1016/j.intell.2009.06.002.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, *47*(5), 667–696 http://doi.org/10.1080/00273171.2012.715555.

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, *95*(2), 129–140 http://doi.org/10.1080/00223891.2012.725437.

Revelle, W. (2015). *Psych: Procedures for Psychological, Psychometric, and Personality Research*. Evanston, Illinois: Northwestern University (Retrieved from http://CRAN.R-project.org/package=psych).

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

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 http://doi.org/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 http://doi.org/10.1016/j.intell.2013.05.002.

Soper, D. S. (2015). *Statistics Calculaters [Software]*. (Available from http://www.danielsoper.com/statcalc).

Sparfeldt, J. R., Buch, S. R., Rost, D. H., & Lehmann, G. (2008). Akkuratesse selbstberichteter Zensuren [The accuracy of self-reported grades in school]. *Psychologie in Erziehung und Unterricht*, *55*, 68–75.

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 http://doi.org/10.1026//0033-3042.50.4.220.

Süß, H. -M. (2001). Prädiktive Validität der intelligenz im schulischen und außerschulischen Bereich [The predictive validity of intelligence in formal and non-formal education]. In E. Stern, & J. Guthke (Eds.), *Perspektiven der Intelligenzforschung* (pp. 109–135). Lengerich: Pabst.

Süß, H. -M., & Beauducel, A. (2011). Intelligenztests und ihre Bezüge zu Intelligenztheorien. [Intelligence tests and their relationships to theories of intelligence]. In L. F. Hornke, M. Amelang, & M. Kersting (Eds.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik*, *Vol. 3*. (pp. 97–234). Göttingen: Hogrefe.

Süß, H. -M., & Beauducel, A. (2015). Modeling the construct validity of the Berlin Intelligence Structure Model. *Estudos de Psicologia (Campinas)*, *32*(1), 13–25 http://doi.org/10.1590/0103-166X2015000100002.

Wallach, D. P. (1997). Learning to control a coal-fired power plant: Empirical results and a model. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics*. *2*. (pp. 82–90). Hampshire: Ashgate Publishers.

Weiß, R. H. (2006). *CFT 20-R - Grundintelligenztest Skala 2 - Revision [Culture Fair Intelligence Test, Scale 2 (Cattell, R.B., 1960) - German revision]*. Göttingen: Hogrefe.

Wilhelm, O. (2005). Measuring reasoning ability. In O. Wilhelm, & R. W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 373–392). London: Sage.

Wirth, J., & Funke, J. (2005). Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme [Dynamic problem solving: Developing and evaluating a new measurement device for controlling complex systems]. In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern: Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie* (pp. 55–72). Wiesbaden: Verlag für Sozialwissenschaften.

Wittmann, W. W., & Hattrup, K. (2004). The relationship between performance in dynamic systems and intelligence. *Systems Research and Behavioral Science*, *21*(4), 393–409 http://doi.org/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: APA.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, *40*(1), 1–14 http://doi.org/10.1016/j.intell.2011.11.003.

Industrial and Organizational Psychology, 8(2), pp 238–268 June 2015.
 doi:10.1017/iop.2015.14

Focal Article

The Assessment of 21st Century Skills in Industrial and Organizational Psychology: Complex and Collaborative Problem Solving

Jonas C. Neubert, Jakob Mainert, André Kretzschmar and Samuel Greiff
University of Luxembourg

In this article, we highlight why and how industrial and organizational psychologists can take advantage of research on 21st century skills and their assessment. We present vital theoretical perspectives, a suitable framework for assessment, and exemplary instruments with a focus on advances in the assessment of human capital. Specifically, complex problem solving (CPS) and collaborative problem solving (ColPS) are two transversal skills (i.e., skills that span multiple domains) that are generally considered critical in the 21st century workplace. The assessment of these skills in education has linked fundamental research with practical applicability and has provided a useful template for workplace assessment. Both CPS and ColPS capture the interaction of individuals with problems that require the active acquisition and application of knowledge in individual or group settings. To ignite a discussion in industrial and organizational psychology, we discuss advances in the assessment of CPS and ColPS and propose ways to move beyond the current state of the art in assessing job-related skills.

When examining the tasks that people perform in their daily workplaces, we see a trend in recent decades toward increases in the importance of nonroutine and interactive tasks. This trend is accompanied by a corresponding decline in routine operations. Jobs that previously entailed repetitive and routine work have been either extended to include nonroutine tasks or removed altogether (e.g., Autor, Levy, & Murnane, 2003; Cascio, 1995). That is, developments in the working world are emphasizing tasks that require active problem solving and that include the need to collaborate with others.

Jonas C. Neubert, Jakob Mainert, André Kretzschmar, and Samuel Greiff, Institute of Cognitive Science and Assessment, University of Luxembourg, Luxembourg.

Jonas Neubert is now at the Brandenburg University of Technology Cottbus-Senftenberg, Germany.

Correspondence concerning this article should be addressed to Samuel Greiff, Institute of Cognitive Science and Assessment, University of Luxembourg, 6 Rue Richard Coudenhove Kalergi, 1359 Luxembourg-Kirchberg, Luxembourg. E-mail: samuel.greiff@uni.lu

By contrast, the number of tasks that people can perform by relying on organizational routines and practices is declining. In economic research, these broader trends have been labeled skill-based technological change, job polarization, and offshoring and have led to a range of insights into the enormous breadth and worldwide scope of the increases in nonroutine and interactive tasks (e.g., Autor, Katz, & Kearney, 2006; Autor et al., 2003; Baumgarten, Geishecker, & Görg, 2010; Becker, Ekholm, & Muendler, 2013; Goos & Manning, 2007; Goos, Manning, & Salomons, 2009; Grossman & Rossi-Hansberg, 2008; Spitz-Oener, 2006).

Widely visible examples of increasing workplace sophistication include the expansion of the role of the modern secretarial staff, the emergence of mechatronics engineers, and the recent extension of the board of directors to include chief operating officers. Secretarial staff members have been taking over former managerial tasks, such as planning, organizing, and supporting meetings and conferences, which is even leading to adaptations of secretarial vocational education. Mechatronics engineering combines several different disciplines into one occupation (i.e., mechanical engineer, electric engineer, and computer scientist), and this job profile itself is an answer to multidisciplinary job demands. On a structural level, organizations increasingly employ chief operating officers, who help to deliver operational excellence in a work environment of increasing complexity. Across industries, chief operating officers have become common on most companies' supervisory boards because of the increase in the numbers of nonroutine and interactive problems. By contrast, only a relatively small number of companies employed executives in this position just 2 decades ago.

As a result of the increasing numbers of nonroutine and interactive tasks, individuals, groups, and organizations are faced with a host of new challenges. Across a wide range of jobs, individuals need to engage in on-the-spot problem-solving behavior without the possibility of resorting to well-defined organizational practices and routines (e.g., Middleton, 2002) and without sufficient time and resources to make decisions about problem-solving measures by following rational models of problem solving (e.g., G. Klein, Orasanu, Calderwood, & Zsambok, 1993; Zsambok & Klein, 1997). In addition, problems increasingly involve the collaboration of multiple individuals from various backgrounds, thus leading to new job requirements; for instance, the integration of diverse pathways toward problem solving within a group comprising members from different backgrounds (e.g., Keane & Nair, 2001; C. Klein, DeRouin, & Salas, 2006, Reiter-Palmon & Illies, 2004).

Organizations consequently need to select, guide, and train individual employees, teams, and leaders who are capable of dealing with emerging job requirements (e.g., Vargas Cortes & Beruvides, 1996). For instance, given the broad range of challenges awaiting a mechatronics engineer taking care of an assembly-line robot in a manufacturing plant, he or she has to be able to anticipate, recognize, and communicate problems within a diverse team. Such an engineer also has to quickly become acquainted with a vast array of complex systems that require immediate and creative solutions when problems occur. In summary, this engineer has to cope with the increased importance of nonroutine and interactive tasks.

As one result, this rise in the importance of nonroutine and interactive tasks has led to broad efforts on multiple levels to specify the accompanying shifts in requirements and skill sets and the facilitation of skills summarized under the umbrella of so-called 21st century skills (e.g., Griffin, McGaw, & Care, 2012; National Research Council, 2012; Organisation for Economic Co-operation and Development 2013c, 2013d). Widely visible, these trends toward nonroutine and interactive tasks have found their way into prominent large-scale assessment efforts such as the Organisation for Economic Co-operation and Development's (OECD's) Programme for the International Student Assessment (PISA; OECD, 2013c, 2013d), which assesses the competencies of more than half a million students worldwide, and the Programme for the International Assessment of Adult Competencies (OECD, 2013b), which targets adult competencies. Whereas those engaged in these efforts used to focus on assessing the skills that individuals acquired during formal education in relation to classical domains such as mathematics and reading, these efforts increasingly feature the assessment of the skills that enable individuals to successfully cope with the requirements of the 21st century and a lifelong perspective.

The realm of general cognitive research has identified two 21st century skills that are strongly related to the demands that have been produced by the changes in the working lives of individuals (e.g., successfully addressing new and complex problems and working collaboratively on a team). The two concepts we deem especially relevant are complex problem solving and collaborative problem solving (CPS and ColPS), a view that is shared by the OECD (OECD, 2013c, 2013d) and other stakeholders (e.g., National Research Council, 2012).

Whereas CPS deals with individuals' transversal skill in successfully handling complex and intransparent situations (i.e., those without a readily apparent solution), requiring the active acquisition and application of knowledge in various domains, ColPS is directed toward problem solving in group settings, adding the necessity of social skills to the ones captured by CPS (e.g., Greiff, 2012; OECD, 2013d). For the mechatronics engineer, these skills can be directly linked to the problems that require attention on a regular basis. Not only do these problems require the gathering of knowledge to generate the understanding of multiple interrelated problem features (e.g., technical and safety requirements, time for implementation, etc.), they also need to be solved in an environment in which vital information is distributed across different members of teams and levels of hierarchies.

Together, CPS and ColPS assess aspects of performance in nonroutine tasks (CPS) and interactive tasks (ColPS) that have been identified as important by Autor et al. (2003) and other researchers (e.g., Cascio, 1995; Spitz-Oener, 2006, see also the literature mentioned above). In addition, researchers studying CPS and ColPS are also committed to conceptual integration and thorough operationalization and assessment, consequently offering solid theoretical and empirical foundations as well as valid assessment methods for their work (e.g., Greiff, Wüstenberg, & Funke, 2012). To this end, both constructs will serve as points of reference for the integration of an assessment of 21st century skills in industrial and organizational (I-O) psychology. In this article, we present CPS and ColPS, their assessment, and potential avenues for the integration of both constructs and their assessment into I-O psychology focusing on skills that enable successful reactions to the challenges of the 21st century.

I-O psychology can be thought of as an applied science with the potential to address, inform, and advise important human-capital (HC) challenges (Cascio & Aguinis, 2008), emphasizing guidance toward practical interventions based on a scientist–practitioner model (Bass, 1974; Dunnette, 1990; Murphy & Saal, 1990; Rupp & Beal, 2007). We believe the field of I-O psychology would benefit from incorporating advances in the definition of 21st century skills and their assessment.

Before taking a closer look at the two constructs of CPS and ColPS, a discussion of three competing approaches to assessment already integrated in I-O psychology is useful because existing assessment methods might in principle allow for the handling of the requirements of the 21st century without the need to resort to new constructs or ways of assessment. Instead of relying on CPS and ColPS, one might argue for the utilization of application-oriented constructs, job-and-work-analysis-based instruments, or well-established constructs targeting basic human functioning. In the following, we take a closer look at all three of these alternatives.

Paths to the Assessment of 21st Century Skills

As a first alternative for employing valid and reliable 21st century skill assessment, we look at constructs that originated from direct observations of the work environment and that developed into nonroutine and interactive tasks. Generally speaking, there are a multitude of constructs addressing the questions of practitioners and business leaders in I-O psychology and management education in an application-oriented way (e.g., building on learning agility: De Meuse, Dai, & Hallenbeck, 2010; Eichinger & Lombardo, 2004; Lombardo & Eichinger, 2000; or the notion of talent and talent management: Cappelli, 2008; Collings & Mellahi, 2009; Ready & Conger, 2007, to give two examples). The constructs and their empirical operationalizations are deeply embedded in their respective fields, and the constructs and operationalizations exhibit close links to practice and application. The point of departure in assessing and using these skills usually occurs when there is an attempt to address a specific need or problem of high visibility and relevance to practitioners and organizations, and the focus of the constructs is consequently related to ways to deal directly with these needs or problems.

As an example, the construct of learning agility originated from the issue of identifying high-potential employees who are capable of performing successfully within a dynamic environment (e.g., Eichinger & Lombardo, 2004; Lombardo & Eichinger, 2000). The relation between learning agility and the trend toward nonroutine and interactive features in the workplace is straightforward: A person confronted with a nonroutine task and the absence of readily applicable routine solutions should directly profit from a higher level of learning agility by being able to better learn from experience in new or first-time conditions (cf. De Meuse et al., 2010).

Hence, selecting individuals on the basis of their learning agility and fostering this agility via training and development seems like a straightforward answer to the trend toward the need to increase performance on nonroutine tasks, thus addressing a highly relevant practical problem (e.g., Dries, Vantilborgh, & Pepermans, 2012). For example, it might be important to select a mechatronics engineer with a highly developed learning agility that allows him or her to actively adapt to changes in the various related domains, such as mechanical and electrical engineering, and that allows him or her to take on future leadership responsibilities. If new safety regulations are introduced, he or she needs to be able to gather the necessary knowledge, assess the influences on various levels, and coordinate the appropriate actions, all of which can be fostered by a high level of learning agility.

On the downside, application-oriented constructs oftentimes lack a clear integration and an explicit connection to their nomological networks. That is, a conceptual and empirical comparison of the commonalities of concepts such as learning agility with regard to well-established and validated constructs and even other application-oriented constructs is largely missing (e.g., for learning agility: Arun, Coyle, & Hauenstein, 2012; DeRue, Ashford, & Myers, 2012; for talent management: Collings & Mellahi, 2009; Lewis & Heckman, 2006; and for a general discussion of the separate discourses oriented toward practice and science: Cohen, 2007; Rynes, Giluk, & Brown, 2007). As an example, whereas learning agility is conceptually related to the notion of reacting to the increasing importance of nonroutine tasks, according to Autor et al. (2003), the conceptual and empirical overlap of such tasks with constructs such as intelligence, personality, or cognitive styles remains unclear (cf. DeRue et al., 2012). In the above example of the mechatronics engineer, it remains unclear whether the successful adaptability of developments in the field can be attributable to a higher level of learning agility or whether such adaptability is a consequence of higher general ability levels, for example, as indicated by intelligence.

At first glance, this does not necessarily lead to problems in situations in which the focus is on the issue of selecting the best fitting mechatronics engineer or in other situations in which the application-oriented construct appropriately meets the environment for which it was developed. However, opportunities to gain more general insights, make long-term predictions, and derive valid conclusions in areas outside the specific focus of the construct have been left unexploited, and theoretical and empirical integration is largely missing. In the example of learning agility, the lack of empirically scrutinized links to personality, intelligence, and other basic constructs leads to doubts about the scientific value of the construct (DeRue et al., 2012) and to missing insights into the range of influenced behaviors and effects on performance. More important, if there is still doubt about whether there is something else other than intelligence in the domain of learning agility, interventions specifically tailored to foster the learning agility of a mechatronics engineer could face massive obstacles (i.e., via boundaries imposed by general levels of intelligence).

Furthermore, and resulting partially from the lack of scientific integration, assessment problems have become widespread in application-oriented constructs, especially with regard to performance measures, jeopardizing the usefulness of the constructs on a fundamental level. When looking at typical instruments that target learning agility, DeRue et al. (2012) identified considerable problems related to both the validity and reliability of the instruments. Lewis and Heckman (2006) showed comparable problems when investigating assessment instruments that targeted the notion of "talent." However, if practitioners and researchers are not able to assess an individual's learning agility in a reliable and valid way, they will not be able to trust practice-oriented advice or conclusions regarding the relations between agility and other constructs and outcomes.

In summary, application-oriented constructs do not necessarily offer the answers required to address the challenges associated with the changes in the working world. Hence, we take a look at a second approach that begins with the very tasks that are becoming less routine and more interactive. Job and work analysis (e.g., Brannick, Levine, & Morgeson, 2007; Fleishman & Reilly, 1992; Rasmussen, Pejtersen, & Goodstein, 1994; Schippmann et al., 2000; Vicente, 1999) offers a long tradition of guidance in matters of personnel selection, training, and planning, going back to the work of Frederick Taylor

(1911) and his *Principles of Scientific Management*, and job and work analysis has been a major part of I-O psychology ever since.

Generally speaking, job and work analysis aims to specify the elements and requirements of a specific job or occupation, usually focused on either the work tasks (task-oriented job analysis, e.g., hierarchical task analysis; Annett & Duncan, 1967; Shepherd, 2001) or the individual performing on the job (worker-oriented job analysis, e.g., position analysis questionnaire; McCormick, Jeanneret, & Mecham, 1972; see, e.g., Clifford, 1994; Dierdorff & Wilson, 2003; Levine, Ash, Hall, & Sistrunk, 1983; Pearlman, 1980, for comparisons between approaches). Building on existing occupations and their descriptions, these efforts have led to considerable knowledge about the elements and requirements of specific jobs and occupations, culminating, for example, in the Occupational Information Network (O*NET) of the United States Department of Labor (http://www.onetonline.org/).

Based on this knowledge of occupations and their associated tasks, job and work analysis also allows for the construction of well-directed assessment instruments targeted precisely at the requirements of a specific job. If researchers and practitioners know what kinds of tasks an individual will most likely be performing in a specific job, they can assemble a corresponding assessment suite to target the associated requirements via job-based tests and work simulations (e.g., Fleishman & Reilly, 1992). In the example of the mechatronics engineer, such instruments may target, for example, the required mathematical skills, knowledge of engineering and technology, and abilities related to deductive reasoning (see also the comprehensive profile of the mechatronics engineer on O*NET, Standard Occupational Classification Code 17-2199.05).

Building on an analysis of job contents and the resulting requirements on the individual level can certainly help to address changes in the working world as it moves toward nonroutine and interactive tasks. Job and work analysis can help researchers and practitioners to quantify the number of changes within jobs and the corresponding requirements (e.g., Cascio, 1995) and can help them to describe, compare, and support newly emerging jobs and work situations within established frames of reference (see, e.g., Naikar, Moylan, & Pearce, 2006; Vicente, 1999). For example, the occupation of mechatronics engineer was specifically added to the O*NET in an effort to include new and emerging occupations in the 21st century (National Center for O*NET Development, 2009), thereby allowing for an analysis of the occupation's tasks and comparisons between this occupation and other (non)engineering jobs.

However, the bottom-up approach of job and work analysis also has its downsides. First, there is the need for a detailed specification of the job and work tasks and the corresponding requirements. There has been incredible progress in this area both in methodology and in content (e.g., the integration of normative and descriptive approaches by Vicente, 1999). Nonetheless, the nonroutine and interactive tasks that are increasing in importance because of the developments in the working world are much harder to grasp and specify specifically because of their nonroutine nature (e.g., Braune & Foshay, 1983; Naikar et al., 2006; Schippmann et al., 2000). Consequently, the effort needed to specify the characteristic tasks in assessment and evaluation is much higher and will increase even more with the ongoing trends toward increases in nonroutine and interactive tasks.

Furthermore, a job-and-work-analysis-based approach to 21st century work environments must rely on a perspective that focuses on already existing and formalized jobs and occupations (even though there are efforts to increase the range to include future developments, e.g., Schneider & Konz, 1989). Because of accelerated technical and social developments and high competitive pressure, there is a vital need to react to changes in work requirements in a proactive way, aligned with the strategic vision and the evolving HC needs of the organization (e.g., Schippmann et al., 2000; but see, e.g., Cascio, 1998; Harvey & Bowin, 1996; and Siddique, 2004, for attempts to integrate job analyses and human resource [HR] strategies). An approach based on job and work analysis necessarily has its limits in this regard because the future jobs and occupations are not yet available for task analyses and job simulations.

Finally, reacting to the increase in the number of nonroutine and interactive tasks based on job and work analysis is a time-consuming and expensive approach (e.g., Levine, Sistrunk, McNutt, & Gael, 1988). Every new job has to be analyzed in detail, and the typical tasks and requirements associated with each job have to be identified. This analysis takes a lot of effort and requires skilled analysts, especially with regard to jobs characterized by nonroutine tasks. This investment can be reduced by efforts to utilize resources, such as the O*NET (e.g., McEntire, Dailey, Osburn, & Mumford, 2006), but such resources are generally not easy to obtain for small and medium-sized organizations, which provide roughly half of all jobs in Western economies (e.g., OECD, 2013a). The great effort necessary to specify the tasks and requirements of a specific job becomes even more problematic when changes in the environment lead to shifting requirements and tasks within these jobs on a regular basis.

The third approach that can be used to assess the challenges resulting from the shifts in the working world toward 21st century skills is to target overarching transversal characteristics that span jobs, problems, and domains by building on established (psychological) research on basic human functioning. The valid and reliable assessment of theoretically well-founded psychological constructs (e.g., intelligence and personality) have historically allowed for extensive progress in I-O psychology (e.g., Barrick & Mount, 1991; Hunter & Schmidt, 1996; Schmidt & Hunter, 1998; Schmidt, Hunter, Outerbridge, & Goff, 1988; Schooler, Mulatu, & Oates, 1999). Furthermore, the constructs are also used as prominent and established markers to assess, analyze, and address HC issues on an individual level (e.g., Barrick & Mount, 1991; Hogan & Holland, 2003; Jones & Schneider, 2006; Seibert & Kraimer, 2001; Weede & Kämpf, 2002). Consequently, these constructs will serve as the third point of departure from which to investigate the consequences of and the answers to the changes in the working world as nonroutine and interactive tasks increase independently of specific jobs or occupations (e.g., Scherbaum, Goldstein, Yusko, Ryan, & Hanges, 2012).

Generally speaking, constructs such as intelligence, personality, or working memory capacity can be viewed as precursors of job performance across a wide array of situations (e.g., Gottfredson, 1997). For example, one would expect that a mechatronics engineer with higher intelligence and a high level of conscientiousness would be generally better at deriving the necessary conclusions in a given problem situation resulting from the challenge of constantly refined and changing technologies (e.g., learning faster and making fewer mistakes when security guidelines related to a new software framework need to be followed).

In contrast to application-oriented constructs, building on constructs such as intelligence leads to theoretically and empirically integrated nomological networks and reliable and valid assessment instruments because the roots of this work are in psychological research. In contrast to a job-and-work-analysis-based approach, the constructs do not have to be bound to specific work tasks or the comprehensive analysis of an occupation to be of considerable use.

Still, given the generality of the constructs and the unspecified interactions between them in concrete situations, personality, intelligence, and other constructs targeting the foundations of human functioning offer only limited help in concrete situations. When looking for individuals who cope well in problem situations that are characterized by a combination of sheer complexity; the need to engage in self-initiated learning behavior (e.g., Warr & Bunce, 1995); and the prerequisite to interact with other individuals, groups, and organizational processes (e.g., Brannick & Prince, 1997), the combination of all potentially relevant constructs makes predictions cumbersome to say the least. Hence, the constructs that dominate current testing and assessment in work organizations as such do not offer the most promising and straightforward solutions when dealing with the changes described by Autor et al. (2003). Awareness of this problem within the realm of I-O research can be seen, for example, in the calls by Brouwers and Van De Vijver (2012) and Oswald and Hough (2012) for the clarification of the pathways between intelligence and actual I-O-related behavior and the calls to take real-life decision processes more seriously in (cognitive) research (e.g., from researchers targeting naturalistic decision making, e.g., G. Klein et al., 1993).

In addition, there are factors that influence the problem situation but are not included in intelligence, personality, and other basic constructs in direct and readily applicable forms. Invoking the mechatronics engineer, we note that he or she needs to analyze, merge, and assemble mechanical, electrical, and electronic components of different generations and within unforeseeably alternating interactions. Furthermore, when production is at full capacity, the targets of production need to be balanced with considerations of the durability of plants and the requirements of working creatively and innovatively with and within research and development teams.

The array of these tasks results in requirements related to (a) the ability to actively generate the information that is needed to see a problem as a complex interplay of developments in the first place, (b) the skills needed to simultaneously balance the changing demands of multiple stakeholder groups, and (c) the prerequisites to interact with various colleagues to pool resources. None of these factors are included in either intelligence assessment or other basic ability tests in a straightforward way (e.g., Funke, 2010). Consequently, targeting HC issues by building on basic constructs and their assessment is restricted, as vital information on performance in situations of rising importance is unavailable. These restrictions are especially relevant when one takes into account the importance of the aforementioned features of decision making in the work environment (G. Klein, 2008; Zsambok & Klein, 1997).

In summary, the comprehensive answer needed to address the call for the adequate assessment of 21st century skills cannot be found in application-oriented constructs situated within I-O psychology or management education, in approaches building on job and work analysis, or in constructs dealing with basic human functioning on a general level. Consequently, we have to look for a direct assessment of 21st century skills as required by the developments in the working environment to be able to build practical advice. CPS and ColPS are considered two prominent representatives of 21st century skills specifically targeting the skills of individuals in problem situations that are characterized by nonroutineness and interactivity. They also allow for an integration of insights across domains and situations, thereby leading to fewer problems in transferring skills across changing environments. Finally, they promote integration into the broader discourse of research and build on solid assessment instruments.

Complex and Collaborative Problem Solving

CPS and ColPS are generally considered integral parts of 21st century skills (Griffin et al., 2012; National Research Council, 2012; OECD, 2013c, 2013d) and have recently been found to have a substantial impact in the area of education. For example, CPS and ColPS are employed in the arguably most important large-scale assessment worldwide, the PISA in its 2012 and 2015 cycles, respectively (OECD, 2013c, 2013d). PISA assesses and compares the skills of students in domains such as mathematics and science across a range of countries to foster policy creation in education, and PISA recently adopted CPS and ColPS as representatives of domain-general transversal skills with clear connections to practice.

Intriguingly enough, whereas both concepts appear to fit the area of I-O psychology rather naturally because they represent the skills necessary to cope with complex and collaborative problems and, hence, the challenges of this context, the widespread application of these concepts in research and practice has mainly been restricted to the field of education. A noticeable exception is the large-scale project "LLLight'in'Europe," in which the CPS skills of more than 4,000 employees of 70 companies in 15 countries are being assessed and analyzed with regard to their relations to income, lifelong learning behaviors, and innovation across various industries, organizations, and jobs (www.lllightineurope.com).

Complex problem solving. CPS targets how humans interact with problems that are characterized by complexity, intransparency, and dynamics, which is sometimes also referred to as dynamic decision making (e.g., Brehmer, 1992; Buchner, 1995; Funke, 2001, 2010; Gonzalez, Lerch, & Lebiere, 2003; Gonzalez, Vanyukov, & Martin, 2005; Schmid, Ragni, Gonzalez, & Funke, 2011). That is, in contrast to historical notions of problem-solving research, CPS targets problem situations featuring a multitude of interrelated elements that have to be actively explored to find a solution, thus requiring the complex interplay of basic cognitive and noncognitive processes (e.g., Fischer, Greiff, & Funke, 2012; Funke, 2010; Osman, 2010).

More specifically, the defining characteristics of problems targeted in CPS are the complexity of the problem structure (i.e., a multitude of interrelated elements), the dynamics of the system (i.e., changes due to time or to interacting with the problem), the interconnectedness of elements (i.e., a change in one part of the system has repercussions in other parts), the multiple goals requiring simultaneous consideration, and the intransparency of the problem situation requiring active investigation (see also the classic definition of complex problems by Buchner, 1995). Naturally, such features are also part of real-life problem solving in the world of I-O psychology in which static problems with a fixed set of options are seldom seen (e.g., Cohen, March, & Olsen, 1972; Smith, 1997). From the perspective of nonroutine tasks as utilized by Autor et al. (2003), there is also a large overlap between the problems targeted in CPS and the larger trends in the working world: Both emphasize the importance of adapting to new situations and problems for which no routine solution is readily available.

With respect to the problem solver, the skills targeted in CPS are clustered around the basic processes of knowledge acquisition and knowledge application (e.g., Fischer et al., 2012; Novick & Bassok, 2005; Osman, 2010). In the example of the mechatronics engineer, knowledge acquisition is related to the gathering of information about a new tool (e.g., an instrument indicating the amount of abrasion), whereas knowledge application can be seen when this knowledge is put to use (e.g., when utilizing the new tool to calibrate a manufacturing robot).

In contrast to basic abilities and constructs (e.g., intelligence or personality), measures of CPS assess performance with a focus on the interaction of individuals in complex problem environments with the individuals' need to actively explore, build, and apply knowledge. In contrast to an approach that builds on job and work analysis, the general importance of CPS constructs is clear and straightforward even without detailed information about the respective jobs or occupations involved. Furthermore and in contrast to application-oriented concepts such as learning agility, CPS is built on a tradition of theoretical and empirical research; thus, it is embedded in a comprehensive nomological network and is measured with reliable and validated assessment instruments.

With regard to this nomological network, CPS has been shown to be conceptually and empirically different from other basic, individual-level constructs such as reasoning ability (Greiff, Fischer, et al., 2013; Sonnleitner, Keller, Martin, & Brunner, 2013; Wüstenberg, Greiff, & Funke, 2012), working memory capacity (Schweizer, Wüstenberg, & Greiff, 2013), and personality as measured by the five-factor model (Greiff & Neubert, 2014). Furthermore, CPS has been shown to be separable from constructs related to specific requirements of the 21st century, such as literacy in information and communication technology, a construct targeting the basic knowledge, skills, and attitudes needed for dealing with computer technology (Greiff, Kretzschmar, Müller, Spinath, & Martin, 2014). With regard to predictive validity, positive and distinct relations between CPS and indicators of successful problem solving in various contexts, ranging from schools and universities to organizations from a range of industries, have been empirically shown on the level of performance measures (e.g., Danner et al., 2011; Greiff, Fischer, et al., 2013; Greiff, Wüstenberg, et al., 2013).

Collaborative problem solving. ColPS, the second construct presented here, is an extension of CPS because it is also related to complex and ill-defined problems. However, whereas CPS targets the skills of individual problems solvers in interacting with complex, intransparent, and dynamic problems, ColPS is dedicated to the assessment of similar skills in interactive settings (i.e., multiple problem solvers working on the same problem; O'Neil, Chuang, & Chung, 2004).

Consequently, processes of knowledge acquisition have to be extended to the group, thus resulting in specific requirements in terms of sharing the understanding and effort required to come to a solution. Also, the pooling of knowledge, skills, and efforts to reach a solution has become vital in both phases of dealing with a problem: knowledge acquisition and application (see, e.g., the definition of ColPS utilized by the OECD in their assessment framework; OECD, 2013d). In the example of the mechatronics engineer introduced before, tasks such as working with experts from other fields to find and address the reasons for the failure of an assembly-line robot clearly involve processes captured by ColPS (e.g., the need to construct a common understanding of the problem at hand).

In line with the rise of collaborative tasks in everyday work environments (e.g., Cascio, 1995; C. Klein et al., 2006), interest in such interactive aspects of problem solving has led to an increase in scientific efforts in recent years (e.g., Greiff, 2012; O'Neil et al., 2004; OECD, 2013d). Still, the nomological network of ColPS and its empirical relations are not as well established as are those for CPS, leaving ample room for future research (e.g., OECD, 2013d). Nonetheless, current research on ColPS has emphasized the connections of ColPS to basic constructs and viable routes for its assessment and application, thereby addressing vital aspects of the problem-solving environments of our times (Greiff, 2012). In light of the rising importance of interactive tasks in the work environment, the skills targeted by ColPS and the assessment of these skills should certainly be incorporated into future discussions in I-O psychology.

The Assessment of Complex and Collaborative Problem Solving

Even in large-scale assessments such as PISA, the assessment of both constructs, CPS and ColPS, has become practical with the help of computer-based microworlds that allow for the simulation of complex and collaborative problems that need to be actively explored and controlled (e.g., Greiff et al., 2012). That is, the skills of individuals in addressing complex and collaborative problems can be assessed directly as those individuals interact with such problems as simulated on a computer or tablet.

To secure the systematic variation in problem features along theoretically derived dimensions (e.g., including a specific type and a specific number of problem features), one usually describes the problems used in these microworlds according to formal frameworks. That is, the computer-based assessment of both CPS and ColPS builds on formal descriptions that enable the construction and systematic variation of problem features in assessment (e.g., a formalization via linear structural equations or finite state automata; Funke, 2001; Greiff & Wüstenberg, 2014).

As a consequence, characteristic features of complex and collaborative problems can be systematically varied in assessment, building on the construct in focus and following theoretically defined and empirically validated dimensions. For example, the number of goals and the number of interconnections that need to be considered can be varied independently. In assessment, these features can be combined systematically, and the characteristics can be compared between problems based on their formalization (e.g., Greiff & Wüstenberg, 2014). By contrast, neither classical tests of intelligence nor job-based simulation instruments can account for these characteristic elements of nonroutine and interactive tasks on the basis of both theoretically sound foundations and reliable and valid instruments.

Building on formally comparable problems and including a specific range of these problem features, the integration of several computer-simulated microworlds in one assessment session leads to the reliable estimation of individual performance levels across specific complex problems (e.g., Greiff et al., 2012). It is important to note that the assessment of CPS and ColPS can tap into actual performance instead of relying on self-reported preferences. The assessment of CPS and ColPS is therefore considerably more closely related to actual performance than are assessment instruments targeting learning agility, cognitive styles, or personality, in which the sole source of information is typically questionnaires filled out by the people themselves.

In addition, because of the use of computer-based assessment, data reflecting the processes of individuals (e.g., when exploring a collaborative problem) become available for analysis. That is, in contrast to assessments via questionnaires as traditionally employed in assessments of basic constructs (e.g., intelligence) or application-oriented concepts (e.g., learning agility), the final performance of individuals can be related to the specific challenges and behavioral foundations of success and failure (e.g., inappropriate exploration strategies for the collaborative problem at hand).

For CPS, existing valid and reliable instruments allow for the estimation of an individual's skills in dealing with complex problems, thereby building a solid foundation for further analyses and interventions in I-O psychology (e.g., MicroDYN: Greiff et al., 2012; GeneticsLab: Sonnleitner et al., 2012; Tailorshop: Putz-Osterloh, 1981; Danner et al., 2011). For ColPS, the development of assessment has progressed tremendously, partly building on the experience already available from CPS assessment. Still, some assessment-related questions need further clarification; for example, how can one take into account multiple problem solvers in one assessment setting (e.g., O'Neil et al., 2004). Nevertheless, we consider it essential to extend problem-solving research to collaborative settings, especially when we keep in mind the rise of related tasks within jobs.

In summary, assessments of CPS and ColPS build on innovative ways for researchers and practitioners to assess the performance of individuals who are dealing with complex and collaborative problems. Because of the computer simulation of problems, requirements found in the example of the mechatronics engineer (e.g., an active search for information when trying to find the source of an error) are systematically included in such assessments and can be systematically varied. Consequently, assessments of CPS and ColPS combine the solid theoretical and psychometric foundations of basic constructs with innovative assessment methods and a focus on application as featured by application-oriented concepts.

Complex and Collaborative Problem Solving and Industrial and Organizational Psychology

Building on the constructs of CPS and ColPS and reliable and valid assessment instruments, how can researchers and practitioners from I-O psychology profit from integrating CPS and ColPS into their toolkits? Answering the call by Cascio and Aguinis (2008) for more HC-related I-O psychology research, we explore the opportunities for and benefits of integrating CPS and ColPS into I-O psychology. These opportunities address the trend in the working world toward nonroutine and interactive tasks as laid out by Autor et al. (2003). To this end, the explorations are grouped around thematic clusters loosely following the classifications of I-O psychology and the field of organizational behavior (e.g., Armstrong, Cools, & Sadler-Smith, 2012; Buchanan & Huczynski, 2010). Against the backdrop of CPS and ColPS, how do individuals enter organizations, strive for career success, develop an actionable transversal skill set, and eventually exceed organizational expectations under the guidance of good leadership?

More specifically, we identify the potential consequences and insights for researchers and practitioners from increasing their attention toward a utilization of CPS and ColPS within I-O psychology. We direct attention to (a) occupational topics further delineated toward personnel selection and career development, (b) human resource (HR) development and learning, and (c) organizational change and the CPS and ColPS side of leadership.

Personnel selection and career development. Generally speaking, both CPS and ColPS are transversal skills that offer researchers the opportunity to better understand general, domain-unspecific problem-solving behaviors that can be explicitly linked to workplace problem solving. Both are promising constructs for bridging the distance between abstract domain-unspecific and general problem-solving skills and concrete work tasks, such as attentive planning, the implementation of complex hardware and software systems (e.g., an assembly line in a milk plant), and the initiation of a system as a team effort. It is important to note that these links can be built on solid foundations in terms of valid and reliable assessment and conceptually clarified constructs, thus reducing the influence of measurement error and conceptual confusion. In the following, we discuss more specifically how personnel selection strategies and career development can profit from a consideration of CPS and ColPS as individual-level prerequisites.

Personnel selection: A matter of fit between candidates and organizations. When entering an organization, it is of great interest for both the new hire and the organization to be compatible with each other in order to pave the way toward successful employment. A fit on multiple organizational levels between an individual's prerequisites and the organization refers to the congruency between the attributes of the person and those of the work environment and encompasses task demands, group phenomena, and organizational features. This so-called person–organization (P-O) fit is of viable interest to selection researchers (Chan, 1996) because certain facets of P-O fit have empirically been shown to predict job-relevant outcomes such as commitment and turnover (e.g., Adkins, Russell, & Werbel, 1994; Chatman, 1989; O'Reilly, Chatman, & Caldwell, 1991; Rynes & Gerhart, 1990).

Most studies have analyzed whether personal values, goals, and interests are congruent with organizational culture, climate, and norms (Adkins et al., 1994; Holland, 1985; O'Reilly et al., 1991; Vancouver & Schmitt, 2006), whereas in general, the cognitive side and, in particular, an ability–demands perspective on P-O fit (e.g., Caplan, 1987; Edwards, 1991; Kristof, 1996) have fallen behind. Hence, it is worthwhile to delineate the contributions of CPS and ColPS to the ability–demands perspective on P-O fit and the implications of these constructs for personnel selection.

Developed as one essential facet of P-O fit (Kirton, 1976; Taylor, 1989), the construct of cognitive misfit, which is operationalized with an emphasis on cognitive styles that range along a continuum from adaption to innovation, invites an exemplary integration of CPS and ColPS. According to Kirton (1976), adaptors solve problems on the basis of incremental change by improving already existing practice, whereas innovators are more likely to initiate change by applying previously unknown ways of doing things. According to empirical results by Chan (1996), cognitive misfit between an individual's cognitive problem-solving style and the demands of the respective work context eventually contributes to increased turnover rates. This style–demands view perceives cognitive styles as unequivocally distinctive from the ability–demands perspective (e.g., Riding, 1997).

A possible synthesis of the cognitive misfit construct with CPS and ColPS and hence the quest for deeper insights into both CPS and ColPS and P-O fit could build on the adaption–innovation continuum and CPS as the construct of choice for an ability–demands perspective. That is, if one combines cognitive styles with CPS, the emerging cognitive-style-skill matrix would highlight the innovator-high-CPS-skill profile as a promising candidate for many 21st century jobs that confront the employee with constantly changing complex problems and require continuous learning of the new. For mechatronics engineers with a drive to innovate, CPS skills and a matching cognitive style are presumably required for creating practical solutions. That is, a mechatronics engineer, or another employee, might fail to solve complex problems in his or her job because the engineer lacks CPS skills even if he or she is equipped with a matching cognitive style and vice versa.

Obviously, the questions of whether employees with a drive to innovate and high CPS skills are readily equipped for 21st century jobs and how they are differentiated from other combinations require further empirical research. However, certainly both CPS and ColPS can be used to increase the efficiency of staff selection procedures and to optimize the degree of fit between potential hires and job roles. Eventually, incorporating these concepts into personnel selection test batteries should add value to companies; for example, by preventing turnovers, which jeopardize HC development intentions and result in losses of organizational knowledge.

Career development: Modern careers and transversal skills. After becoming part of an organization on the basis of mutual compatibility, an employee's attention usually centers on the potential to grow personally and to ascend the career ladder. Turning to the individual confronted with the changes in the working world, the reduced significance of traditional organizational career paths that rely on organizational structures with a paternalistic approach to career management, vertical mobility, and reasonable stability becomes a central factor (e.g., Allred, Snow, & Miles, 1996; Arthur & Rousseau, 1996).

In modern organizational work environments, the classic career approach has broadly been replaced by new career paradigms that deemphasize organizational factors and stress the importance of the individual and his or her skill set (Arthur & Rousseau, 1996; Bird, 1994; Greenhaus, Callanan, & Kaplan, 1995; Hall & Mirvis, 1996). Some of these newer concepts speak, for instance, of a protean or boundaryless career, which emphasizes the transferability of skills and acknowledges that individuals need to take responsibility for managing their own careers, to reveal lateral mobility, and to take on different roles in multiple projects.

Whereas self-knowledge, interpersonal knowledge, and environmental knowledge are identified as the key factors in the literature on new careers (e.g., Anakwe, Hall, & Schor, 2000), the requirements that individuals have to face also show considerable overlap with the skills targeted by CPS and ColPS. Both research directions emphasize the central role of transversal skills and transferable knowledge at work (Anakwe et al., 2000; Griffin et al., 2012).

Reacting successfully to new situations that cannot be handled solely on the basis of factual knowledge and experience in a fixed environment necessarily builds on transferable domain-generalizable skills such as CPS and ColPS. Because of the transversal nature of both skills and in line with developments in modern careers, benefits from developing one's CPS and ColPS skills will not be restricted to a specific area of application (e.g., a specific tool, job, or organization). For example, mechatronics engineers with an enhanced CPS and ColPS skill set will be able to utilize these skills in several different work environments, for instance, when they are transferring a project from the pilot to the production phase. These mechatronics engineers will also be able to translate their skills into palpable career development even if they change from one organization or job role to another, due to the rather context-independent nature of the skills. In short, CPS and ColPS are both factors with incredible potential for selecting, developing, and supporting individuals in their modern day careers and spearheading the way toward tangible insights and evidence-based reactions to changes in career development.

Human resource development and learning. Any career requires the development of an actionable skill set for heightened task performance. HR practitioners are concerned with the skill development and lifelong learning of their organization's employees and definitely require access to pertinent high-quality information. The availability of transversal skills, such as CPS and ColPS, opens gateways for improvements in HR practice, which, at the moment, mostly relies on the development of domain-specific knowledge (e.g., via job and work analysis; e.g., Cascio, 1998; Levine et al., 1988). Thus, the question arises: How can transversal skills such as CPS and ColPS be integrated into occupational assessment and trainings?

An example might be a product designer in the research and development department of a company in the electronics industry. Fulfilling a variety of individual and team tasks, including the handling of complex and innovative products, the designer has a job that requires CPS and ColPS skills to a great extent because domain-specific knowledge is not sufficient for dealing with the changing requirements. Connections to specific competencies are required in that the designer has to ensure the aesthetic quality of interfaces and the alignment with product guidelines; it is his or her duty to conduct tests of prototypes and existing products, thus the job requires extensive experience in the domain as well as problem-solving skills that allow the designer to react to unforeseen challenges.

Although building on the foundation of specific knowledge, the designer, within most of these activities, continuously acquires and applies knowledge and collaborates on interdisciplinary teams with other designers, engineers, and industrial psychologists. CPS and ColPS allow for the targeting of skills required to successfully deal with situations when prior knowledge and experience are either scarce or not sufficient for dealing with problems on a routine level. As a consequence, they are a valuable tool for HR departments in assessing their employees' HC beyond measures of education or experience, such as tenure. Further, CPS and ColPS are considered to be indispensable assets when one is acquiring and consolidating specific knowledge and experience in terms of lifelong learning (OECD, 2013c). This initial picture of how CPS and ColPS might benefit HR practices can serve only as a point of departure, and specific connections and insights into the interplay of the multitude of related factors are a rich field of inquiry for both future research and practical application.

Organizational change and the complex and collaborative problem-solving side of leadership. It is unlikely that solutions to nonroutine problems will be accomplished without a large degree of support from organizations. Representing one pathway by which this organizational support can be provided, leadership is thought to be a critical resource for developing appropriate problem-solving skills (Reiter-Palmon & Illies, 2004). In fact, Reiter-Palmon and Illies (2004) suggested avenues by which organizational leaders can facilitate early stage cognitive processes in an effort to enhance the problem solving of their employees. For instance, leaders would do well to encourage their subordinates to take more time, communicate with their team, and regard different perspectives for the definition and construction of a problem. Taking a similar approach to examining the possibilities and effects of facilitating the CPS and ColPS of employees through various organizational processes (e.g., leadership) certainly promises to lead to interesting, practical, and highly relevant insights.

Shifting the focus from the leader's organizational role in facilitating the problem solving of subordinates to the leader's own problem-solving ability, we invite leadership researchers to follow our emphasis on the cognitive side of leadership, where CPS and ColPS could spice up existing research directions. In times of organizational change, leaders have to be prepared for the unexpected to be able to provide adequate guidance to their subordinates. With their emphasis on knowledge acquisition, knowledge application, and shared understanding in groups, CPS and ColPS contribute to a comprehensive assessment of leadership skills. Well-established theories of organizational leadership employ a range of classic qualitative measures, including the Multifactor Leadership Questionnaire (Bass & Avolio, 1990) or the Leader Opinion Questionnaire (Fleishman, 1989), to shed light on a leader's interaction style with subordinates.

Several authors have argued that the focus of research on organizational leadership and assessment ought to incorporate the (cognitive) substance of leadership and should not focus exclusively on leadership styles (e.g., Day & Lord, 1988; Jacobs & Jaques, 1987). This view suggests that organizational leadership should be perceived as a form of skilled performance grounded in the leader's ability to solve complex and ill-defined organizational problems (Mumford, Mobley, Reiter-Palmon, Uhlman, & Doares, 1991; Mumford, Zaccaro, Harding, Jacobs, & Fleishman, 2000; Zaccaro et al., 1997).

Apparently, qualitative leadership style measures do not account for the cognitive side of leadership, and hence, different leader assessment strategies involving tools explicitly designed to assess CPS skills are required. For instance, CPS subskills such as problem construction, information encoding, and solution implementation and monitoring skills (e.g., Fischer et al., 2012) have been identified by Mumford et al. (1991) as important parts of leadership purely on the cognitive side. Further, Zaccaro, Mumford, Connelly, Marks, and Gilbert (2000) extended leaders' problem solving to collaborative aspects by acknowledging that leadership is embedded in a social context and by emphasizing social skills that reflect an understanding of people and social systems, especially during organizational change.

Whereas these authors used complex organizational scenarios with open questions in a paper-and-pencil format that relied on the raters' own interpretations, Marshall-Mies et al. (2000) had already introduced an online computer-based leader assessment strategy by building on a predetermined set of choices across comparable scenarios. However, neither approach can account for the procedural, dynamic, and complexity-related aspects of interacting with leadership problems.

By contrast, microworlds as exploited by Greiff and colleagues (Greiff & Wüstenberg, 2014; Greiff et al. 2012) in educational settings are able to account for such aspects by requiring the active gathering, integration, and application of knowledge. Consequently, the measures can be seen as reliable proxies for the assessment of at least some of the prerequisites for successful leadership, incorporating the skills to acquire and apply knowledge across situations and problem solving in groups. The CPS and ColPS abilities of leaders have been identified as key factors in producing organizational transformations (e.g., Bruch, Spychala, & Wiegel, 2013). With the help of CPS and ColPS assessment, researchers and practitioners have the tools available to incorporate these abilities into empirical research and practical application.

Conclusion

Coming back to the notion of a world of increasing complexity and the increasing importance of nonroutine and interactive tasks in the daily working lives of individuals, we must ask whether the proposed integration of CPS and ColPS will be able to overcome the lack of satisfactory approaches to the assessment of 21st century skills in I-O psychology. In contrast to (a) basic psychological constructs, which are missing connections to application; (b) job-and-work-analysis-based approaches, which require huge efforts specific to a job or occupation; and (c) application-oriented constructs, which suffer from a lack of integration and viable ways of assessment, we indeed believe that CPS and ColPS have the potential to serve as a point of departure for future development.

Clearly, if we as researchers and practitioners do not strive for an integration of the various components that influence the interaction of individuals and groups with problem situations characterized by complexity and dynamic and interactive features, we may as well stick with a separate assessment of different basic constructs and analyze their influences separately on an abstract level. If we are not interested in developing an encompassing perspective that integrates concepts beyond our specific line of inquiry, and if we do not mind problems of assessment hampering scientific and applied progress, we may also be satisfied with application-oriented constructs without a path to theoretical integration and sound empirical application. If we look at highly standardized production jobs, stable work environments, or occupations that will remain the same across the years, we may be inclined to continue to put our efforts toward job and work analysis.

However, if researchers and practitioners pursue the route of bringing I-O psychology forward, both theoretically and practically, CPS and ColPS are two lines of inquiry worth considering. Combining a focus on ill-defined and complex problems with rigorous empirical research and clear paths for application, these constructs naturally fit the developments of the 21st century and provide the necessary components for the field of I-O psychology. We believe the field of I-O psychology would benefit from practitioners and researchers incorporating advances in the definitions of these 21st century skills and the tools used to assess these skills that are currently part of the OECD and PISA programs. Furthermore, an incorporation of these advances would also facilitate international collaboration between I-O psychologists in the United States and researchers and practitioners in Europe and Asia, where the PISA assessments and OECD programs are already having a visible impact on educational and school-to-work policies.

Following the example of the PISA assessments and the OECD, including CPS and ColPS in the arena of I-O psychology might lead to interesting developments: CPS and ColPS might offer ways to preserve the benefits of classical constructs (e.g., intelligence) in terms of well-connected nomological networks, empirically scrutinized links to the broader scientific context, and reliable and valid assessment instruments and might offer ways to combine these benefits with clear routes for application and practice-oriented advice. Ultimately, this could lead to the replacement of assessment instruments that target basic constructs with instruments that are more closely aligned with the developments in the working world of the 21st century. As the time and money available for assessment is almost always limited, a direct assessment of 21st century skills could actually be the preferable option compared with basic constructs that lack clear connections to practice. Integrating insights from a range of classical domains while acknowledging the fact that the complex problems we encounter on an everyday basis need more than the sum of basic (non)cognitive processes (cf. Funke, 2010) also raises questions related to the overcoming of the scientist–practitioner gap (Cascio & Aguinis, 2008).

Analyzing current trends in practice and research, Cascio and Aguinis (2008) indicated a serious disconnect between the knowledge that I-O psychologists are producing and the knowledge that practitioners are consuming (e.g., Rynes, Colbert, & Brown, 2002). Consequently, a scientist–practitioner gap persists (Aguinis & Pierce, 2008; Anderson, 2007), and I-O psychology has yet failed to provide sufficient answers to public policy or to management practices about HC trends in a changing world of work in the 21st century. The current state of affairs is certainly something with which we cannot be satisfied, and the examples we presented of application-oriented constructs clearly indicate the drawbacks of discourses that do not manage to connect scientific rigor and practical relevance for both practitioners and researchers alike. Constructing a bridge across this gap might indeed build on the insights presented here for the direct assessment of practically relevant 21st century skills such as CPS and ColPS. We are eager to find out whether application-oriented researchers and colleagues dedicated to classical constructs (e.g., intelligence) agree with this opinion.

We are unsure whether the detailed analyses of requirements resulting from task-specific characteristics as identified by job and work analysis might be a good starting point for such endeavors, but the question remains as to whether I-O psychology really needs to base each and every instance of assessment and intervention on such costly grounds. If we can identify larger trends, such as the ones identified for nonroutine and interactive tasks, we might actually be better off with broader categories of skills that are applicable across specific contexts and jobs. To be fair, domain-specific knowledge and tenure will always be relevant predictors of job performance in specific situations, but the question remains as to whether I-O psychologists, HR professionals, and organizations can take the development of these factors into account and base their reasoning on broader 21st century skills that make detailed analyses of narrow job tasks superfluous.

We discuss approaches currently available to I-O researchers and practitioners, highlighting their benefits and drawbacks in terms of theoretical, empirical, and practical integration, and present an alternative: a direct assessment of 21st century skills such as CPS and ColPS. Tushman and O'Reilly (2007) reasoned that the gap between actual research and practical concerns in I-O psychology reduces the impact of the field's research and undermines the external validity of its theories. Even further, Anderson, Herriot, and Hodgkinson (2001) perceived the separation process between academics and practitioners as a threat to the core values of the discipline, which as Anderson et al. (2001) warned could seriously impede the field. To advance the discussion in I-O psychology, we presented CPS and ColPS, two 21st century skills that might provide an answer for overcoming the scientist–practitioner gap with their theoretical line of inquiry, their focus on reliable and valid assessment, and their routes to application in I-O psychology. Furthermore, as the examples from several areas of application in I-O psychology should have made clear, the possibilities of combining CPS and ColPS with domain-specific questions in I-O psychology—targeting, for example, HR development and leadership—offer promising opportunities for application-oriented conclusions and interventions as well as further research and scientific inquiry.

In closing, CPS and ColPS are two approaches to 21st century skill assessment that will contribute to the development and prosperity of I-O psychology by combining the strengths of basic and well-validated constructs with the richness and palpable value of application-oriented concepts for progress in insights and innovation. Let's get started!

References

Adkins, C. L., Russell, C. J., & Werbel, J. D. (1994). Judgments of fit in the selection process: The role of work value congruence. *Personnel Psychology*, *47*, 605–623. doi:10.1111/j.1744-6570.1994.tb01740.x

Aguinis, H., & Pierce, C. A. (2008). Enhancing the relevance of organizational behavior by embracing performance management research. *Journal of Organizational Behavior*, *29*, 139–145. doi:10.1002/job.493

Allred, B. B., Snow, C. C., & Miles, R. E. (1996). Characteristics of managerial careers in the 21st century. *Academy of Management Executive*, *10*(4), 17–27. doi:10.5465/AME.1996.3145316

Anakwe, U. P., Hall, J. C., & Schor, S. M. (2000). Knowledge-related skills and effective career management. *International Journal of Manpower*, *21*, 566–579. doi:10.1108/01437720010379024

Anderson, N. (2007). The practitioner-researcher divide revisited: Strategic-level bridges and the roles of IWO psychologists. *Journal of Occupational and Organizational Psychology*, *80*, 175–183. doi:10.1348/096317907´187237

Anderson, N., Herriot, P., & Hodgkinson, G. P. (2001). The practitioner-researcher divide in industrial, work and organizational (IWO) psychology: Where are we now, and where do we go from here? *Journal of Occupational and Organizational Psychology*, *74*, 391–411. doi:10.1348/096317901167451

Annett, J., & Duncan, K. (1967). Task analysis and training design. *Occupational Psychology*, *41*, 211–221.

Armstrong, S. J., Cools, E., & Sadler-Smith, E. (2012). Role of cognitive styles in business and management: Reviewing 40 years of research: Role of cognitive styles in business and management. *International Journal of Management Reviews*, *14*, 238–262. doi:10.1111/j.1468-2370.2011.00315.x

Arthur, M. B., & Rousseau, D. M. (1996). A career lexicon for the 21st century. *Academy of Management Perspectives*, *10*(4), 28–39. doi:10.5465/AME.1996.3145317

Arun, N., Coyle, P. T., & Hauenstein, N. (2012). Learning agility: Still searching for clarity on a confounded construct. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *5*, 290–293. doi:10.1111/j.1754-9434.2012.01447.x

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The polarization of the U.S. labor market. *American Economic Review*, *96*(2), 189–194. doi:10.1257/000282806777212620

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*, 1279–1333. doi:10.1162/003355303322552801

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*, 1–26. doi:10.1111/j.1744-6570.1991.tb00688.x

Bass, B. M. (1974). The substance and the shadow. *American Psychologist*, *29*, 870–886. doi:10.1037/h0037515

Bass, B. M., & Avolio, B. J. (1990). *Transformational leadership development: Manual for the multifactor leadership questionnaire*. Palo Alto, CA: Consulting Psychologists Press.

Baumgarten, D., Geishecker, I., & Görg, H. (2010). Offshoring, tasks, and the skill-wage pattern. *CEGE Discussion Paper*, *98*. Retrieved from http://ssrn.com/abstract=1574784. doi:10.2139/ssrn.1574784

Becker, S. O., Ekholm, K., & Muendler, M.-A. (2013). Offshoring and the onshore composition of tasks and skills. *Journal of International Economics*, *90*, 91–106. doi:10.1016/j.jinteco.2012.10.005

Bird, A. (1994). Careers as repositories of knowledge: A new perspective on boundaryless careers. *Journal of Organizational Behavior*, *15*, 325–344. doi:10.1002/job.4030150404

Brannick, M. T., Levine, E. L., & Morgeson, F. P. (2007). *Job and work analysis: Methods, research, and applications for human resource management* (2nd ed.). Los Angeles, CA: Sage.

Brannick, M. T., & Prince, C. (1997). An overview of team performance measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 3–16). Mahwah, NJ: Erlbaum.

Braune, R., & Foshay, W. R. (1983). Towards a practical model of cognitive/information processing task analysis and schema acquisition for complex problem-solving situations. *Instructional Science*, *12*, 121–145.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, *81*, 211–241.

Brouwers, S. A., & Van De Vijver, F. J. R. (2012). Intelligence 2.0 in I–O psychology: Revival or contextualization? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *5*, 158–160. doi:10.1111/j.1754-9434.2012.01422.x

Bruch, H., Spychala, A., & Wiegel, J. L. (2013). Wege zur hochleistungsorganisation: Wie die Graubündner Kantonalbank energie mobilisiert und erhält [Pathways to a high performance organization: How the Graubündner Kantonalbank mobilizes and maintains energy]. *Personalführung*, *2013*(2), 20–27.

Buchanan, D. A., & Huczynski, A. (2010). *Organizational behaviour* (7th ed.). Harlow, United Kingdom: Pearson Education.

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* (pp. 27–63). Hillsdale, NJ: Erlbaum.

Caplan, R. D. (1987). Person–environment fit theory and organizations: Commensurate dimensions, time perspectives, and mechanisms. *Journal of Vocational Behavior*, *31*, 248–267. doi:10.1016/0001-8791(87)90042-X

Cappelli, P. (2008). Talent management for the twenty-first century. *Harvard Business Review*, *86*(3), 74–81.

Cascio, W. F. (1995). Whither industrial and organizational psychology in a changing world of work? *American Psychologist*, *50*, 928–939. doi:10.1037/0003-066X.50.11.928

Cascio, W. F. (1998). *Managing human resources: Productivity, quality of work life, profits*. Boston, MA: Irwin McGraw-Hill.

Cascio, W. F., & Aguinis, H. (2008). Research in industrial and organizational psychology from 1963 to 2007: Changes, choices, and trends. *Journal of Applied Psychology*, *93*, 1062–1081. doi:10.1037/0021-9010.93.5.1062

Chan, D. (1996). Cognitive misfit of problem-solving style at work: A facet of person-organization fit. *Organizational Behavior and Human Decision Processes*, *68*, 194–207. doi:10.1006/obhd.1996.0099

Chatman, J. A. (1989). Improving interactional organizational research: A Model of person-organization fit. *Academy of Management Review*, *14*, 333–349. doi:10.5465/AMR.1989.4279063

Clifford, J. P. (1994). Job analysis: Why do it, and how should it be done? *Public Personnel Management*, *23*, 321–340. doi:10.1177/009102609402300211

Cohen, D. J. (2007). The very separate worlds of academic and practitioner publications in human resource management: Reasons for the divide and concrete solutions for bridging the gap. *Academy of Management Journal*, *50*, 1013–1019. doi:10.5465/AMJ.2007.27151946

Cohen, M. D., March, J. G., & Olsen, J. P. (1972). "A garbage can model of organizational choice." *Administrative Science Quarterly*, *17*, 1–25.

Collings, D. G., & Mellahi, K. (2009). Strategic talent management: A review and research agenda. *Human Resource Management Review*, *19*, 304–313. doi:10.1016/j.hrmr.2009.04.001

Danner, D., Hagemann, D., Holt, D. V., Hager, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). Measuring performance in dynamic decision making. *Journal of Individual Differences*, *32*, 225–233.

Day, D. V., & Lord, R. G. (1988). Executive leadership and organizational performance: Suggestions for a new theory and methodology. *Journal of Management*, *14*, 453–464. doi:10.1177/014920638801400308

De Meuse, K. P., Dai, G., & Hallenbeck, G. S. (2010). Learning agility: A construct whose time has come. *Consulting Psychology Journal: Practice and Research*, *62*, 119–130. doi:10.1037/a0019988

DeRue, D. S., Ashford, S. J., & Myers, C. G. (2012). Learning agility: In search of conceptual clarity and theoretical grounding. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *5*, 258–279. doi:10.1111/j.1754-9434.2012.01444.x

Dierdorff, E. C., & Wilson, M. A. (2003). A meta-analysis of job analysis reliability. *Journal of Applied Psychology*, *88*, 635–646. doi:10.1037/0021-9010.88.4.635

Dries, N., Vantilborgh, T., & Pepermans, R. (2012). The role of learning agility and career variety in the identification and development of high potential employees. *Personnel Review*, *41*, 340–358. doi:10.1108/00483481211212977

Dunnette, M. D. (1990). Blending the science and practice of industrial and organizational psychology: Where are we and where are we going? In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1, pp. 1–27). Palo Alto, CA: Consulting Psychologists Press.

Edwards, J. R. (1991). Person-job fit: A conceptual integration, literature review, and methodological critique. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial/Organizational Psychology* (Vol. 9, pp. 383–357). London, United Kingdom: Wiley.

Eichinger, R. W., & Lombardo, M. M. (2004). Learning agility as a prime indicator of potential. *Human Resource Planning*, *27*(4), 12–15.

Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, *4*(1), 19–42. doi:10.7771/1932-6246.1118

Fleishman, E. A. (1989). *Manual for the Leadership Opinion Questionnaire* (Rev. ed.). Rosemont, IL: National Computer Systems.

Fleishman, E. A., & Reilly, M. E. (1992). *Handbook of human abilities: Definitions, measurements, and job task requirements.* Palo Alto, CA: Consulting Psychologists Press.

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

Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*, 133–142. doi:10.1007/s10339-009-0345-0

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, *27*, 591–635.

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

Goos, M., & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *Review of Economics and Statistics*, *89*, 118–133. doi:10.1162/rest.89.1.118

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. *The American Economic Review*, *99*(2), 58–63.

Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence*, *24*, 79–132. doi:10.1016/S0160-2896(97)90014-3

Greenhaus, J. H., Callanan, G. A., & Kaplan, E. (1995). The role of goal setting in career management. *International Journal of Career Management*, *7*(5), 3–12. doi:10.1108/09556219510093285

Greiff, S. (2012). From interactive to collaborative problem solving: Current issues in the Programme for International Student Assessment. *Review of Psychology*, *19*, 111–121.

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence*, *41*, 579–596. doi:10.1016/j.intell.2013.07.012

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

Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, *36*, 37–48. doi:10.1016/j.lindif.2014.08.003

Greiff, S., & Wüstenberg, S. (2014). Assessment with microworlds using MicroDYN: Measurement invariance and latent mean comparisons: Psychometric properties across several student samples and blue-collar workers. *European Journal of Psychological Assessment*, *30*, 304–314. doi:10.1027/1015-5759/a000194

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

Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. New York, NY: Springer.

Grossman, G. M., & Rossi-Hansberg, E. (2008). Trading tasks: A simple theory of offshoring. *American Economic Review*, *98*(5), 1978–1997. doi:10.1257/aer.98.5.1978

Hall, D. T., & Mirvis, P. H. (1996). The new protean career. In D. T. Hall (Ed.), *The career is dead—Long live the career: A relational approach to careers* (pp. 15–45). San Francisco, CA: Jossey-Bass.

Harvey, D. F., & Bowin, R. B. (1996). *Human resource management: An experiential approach*. Englewood Cliffs, NJ: Prentice-Hall.

Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology*, *88*, 100–112. doi:10.1037/0021-9010.88.1.100

Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments*. Odessa, FL: Psychological Assessment Resources.

Hunter, J. E., & Schmidt, F. L. (1996). Intelligence and job performance: Economic and social implications. *Psychology, Public Policy, and Law*, *2*, 447–472. doi:10.1037/1076-8971.2.3-4.447

Jacobs, T. O., & Jaques, E. (1987). Leadership in complex systems. In J. Zeidner (Ed.), *Human productivity enhancement* (Vols. 1–2, pp. 7–65). New York, NY: Praeger.

Jones, G., & Schneider, W. J. (2006). Intelligence, human capital, and economic growth: A Bayesian averaging of classical estimates (BACE) approach. *Journal of Economic Growth*, *11*, 71–93. doi:10.1007/s10887-006-7407-2

Keane, A. J., & Nair, P. B. (2001). Problem solving environments in aerospace design. *Advances in Engineering Software*, *32*, 477–487. doi:10.1016/S0965-9978(00)00108-3

Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, *61*, 622–629. doi:10.1037/0021-9010.61.5.622

Klein, C., DeRouin, R. E., & Salas, E. (2006). Uncovering workplace interpersonal skills: A review, framework, and research agenda. In G. P. Hodgkinson & J. K. Ford (Eds.), *International review of industrial and organizational psychology* (pp. 79–126). West Sussex, England: Wiley.

Klein, G. (2008). Naturalistic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*, 456–460. doi:10.1518/001872008´288385

Klein, G., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Kristof, A. L. (1996). Person–organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, *49*, 1–49. doi:10.1111/j.1744-6570.1996.tb01790.x

Levine, E. L., Ash, R. A., Hall, H., & Sistrunk, F. (1983). Evaluation of job analysis methods by experienced job analysts. *The Academy of Management Journal*, *26*, 339–348. doi:10.2307/255981

Levine, E. L., Sistrunk, F., McNutt, K. J., & Gael, S. (1988). Exemplary job analysis systems in selected organizations: A description of process and outcomes. *Journal of Business and Psychology*, *3*, 3–21. doi:10.1007/BF01016745

Lewis, R. E., & Heckman, R. J. (2006). Talent management: A critical review. *Human Resource Management Review*, *16*, 139–154. doi:10.1016/j.hrmr.2006.03.001

Lombardo, M. M., & Eichinger, R. W. (2000). High potentials as high learners. *Human Resource Management*, *39*, 321–329. doi:10.1002/1099-050X(200024)39:4<321::AID-HRM4>3.0.CO;2-1

Marshall-Mies, J. C., Fleishman, E. A., Martin, J. A., Zaccaro, S. J., Baughman, W. A., & McGee, M. L. (2000). Development and evaluation of cognitive and metacognitive measures for predicting leadership potential. *The Leadership Quarterly*, *11*, 135–153. doi:10.1016/S1048-9843(99)00046-6

McCormick, E. J., Jeanneret, P. R., & Mecham, R. C. (1972). A study of job characteristics and job dimensions as based on the Position Analysis Questionnaire (PAQ). *Journal of Applied Psychology*, *56*, 347–368. doi:10.1037/h0033099

McEntire, L. E., Dailey, L. R., Osburn, H. K., & Mumford, M. D. (2006). Innovations in job analysis: Development and application of metrics to analyze job data. *Human Resource Management Review*, *16*, 310–323. doi:10.1016/j.hrmr.2006.05.004

Middleton, H. (2002). Complex problem solving in a workplace setting. *International Journal of Educational Research*, *37*, 67–84. doi:10.1016/S0883-0355(02)00022-8

Mumford, M. D., Mobley, M. I., Reiter-Palmon, R., Uhlman, C. E., & Doares, L. M. (1991). Process analytic models of creative capacities. *Creativity Research Journal*, *4*, 91–122. doi:10.1080/10400419109534380

Mumford, M. D., Zaccaro, S. J., Harding, F. D., Jacobs, T. O., & Fleishman, E. A. (2000). Leadership skills for a changing world: Solving complex social problems. *The Leadership Quarterly*, *11*, 11–35. doi:10.1016/S1048-9843(99)00041-7

Murphy, K. R., & Saal, F. E. (1990). What should we expect from scientist–practitioners? In K. R. Murphy & F. E. Saal (Eds.), *Psychology in organizations: Integrating science and practice* (pp. 49–66). Hillsdale, NJ: Erlbaum.

Naikar, N., Moylan, A., & Pearce, B. (2006). Analysing activity in complex systems with cognitive work analysis: Concepts, guidelines and case study for control task analysis. *Theoretical Issues in Ergonomics Science*, *7*, 371–394. doi:10.1080/14639220500098821

National Center for O*NET Development. (2009). *New and emerging occupations of the 21st century: Updating the O*NET-SOC taxonomy* [O*NET Reports and Documents]. Raleigh, NC: Author.

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

Novick, L. R., & Bassok, M. (2005). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 321–349). New York, NY: Cambridge University Press.

O'Neil, H. F., Chuang, S., & Chung, G. K. W. K. (2004). *Issues in the computer-based assessment of collaborative problem solving* (CSE Report No. 620). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

O'Reilly, C. A., Chatman, J. A., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person–organization fit. *Academy of Management Journal*, *34*, 487–516. doi:10.2307/256404

Organisation for Economic Co-operation and Development. (2013a). Employment by size class. In *Entrepreneurship at a glance 2013*. Paris, France: Author. Retrieved from http://www.keepeek.com/Digital-Asset-Management/oecd/industry-and-services/entrepreneurship-at-a-glance-2013/employment-by-size-class_entrepreneur_aag-2013-8-en#page1

Organisation for Economic Co-operation and Development. (2013b). *OECD skills outlook 2013: First results from the survey of adult skills*. Paris, France: Author. doi:10.1787/9789264204256-en

Organisation for Economic Co-operation and Development. (2013c). *PISA 2012 assessment and analytical framework*. Paris, France: Author.

Organisation for Economic Co-operation and Development. (2013d). *PISA 2015 draft collaborative problem solving framework*. Paris, France: Author. Retrieved from http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf

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

Oswald, F. L., & Hough, L. (2012). I–O 2.0 From intelligence 1.5: Staying (just) behind the cutting edge of intelligence theories. *Industrial and Organizational Psychology*, *5*, 172–175. doi:10.1111/j.1754-9434.2012.01425.x

Pearlman, K. (1980). Job families: A review and discussion of their implications for personnel selection. *Psychological Bulletin*, *87*, 1–28. doi:10.1037/0033-2909.87.1.1

Putz-Osterloh, W. (1981). Über die beziehung zwischen Testintelligenz und problemlöseerfolg [The relation between test intelligence and problem solving success]. *Zeitschrift Für Psychologie Mit Zeitschrift Für Angewandte Psychologie*, *189*(1), 79–100.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. New York, NY: Wiley.

Ready, D. A., & Conger, J. A. (2007). Make your company a talent factory. *Harvard Business Review*, *85*(6), 68–77.

Reiter-Palmon, R., & Illies, J. J. (2004). Leadership and creativity: Understanding leadership from a creative problem-solving perspective. *The Leadership Quarterly*, *15*, 55–77. doi:10.1016/j.leaqua.2003.12.005

Riding, R. J. (1997). On the nature of cognitive style. *Educational Psychology*, *17*, 29–49. doi:10.1080/0144341970170102

Rupp, D. E., & Beal, D. (2007, July). Checking in with the scientist–practitioner model: How are we doing? *The Industrial-Organizational Psychologist*, *45*(1), 35–40.

Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, *41*, 149–174. doi:10.1002/hrm.10029

Rynes, S. L., & Gerhart, B. (1990). Interviewer assessments of applicant "fit": An exploratory investigation. *Personnel Psychology*, *43*, 13–35. doi:10.1111/j.1744-6570.1990.tb02004.x

Rynes, S. L., Giluk, T. L., & Brown, K. G. (2007). The very separate worlds of academic and practitioner periodicals in human resource management: Implications for evidence-based management. *Academy of Management Journal*, *50*, 987–1008. doi:10.5465/AMJ.2007.27151939

Scherbaum, C. A., Goldstein, H. W., Yusko, K. P., Ryan, R., & Hanges, P. J. (2012). Intelligence 2.0: Reestablishing a research program on *g* in I–O psychology. *Industrial and Organizational Psychology*, *5*, 128–148. doi:10.1111/j.1754-9434.2012.01419.x

Schippmann, J. S., Ash, R. A., Battista, M., Carr, L., Eyde, L. D., Hesketh, B., . . . Sanchez, J. I. (2000). The Practice of Competency Modeling. *Personnel Psychology*, *53*, 703–740.

Schmid, U., Ragni, M., Gonzalez, C., & Funke, J. (2011). The challenge of complexity for cognitive systems. *Cognitive Systems Research*, *12*, 211–218. doi:10.1016/j.cogsys.2010.12.007

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262–274. doi:10.1037/0033-2909.124.2.262

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology*, *73*, 46–57. doi:10.1037/0021-9010.73.1.46

Schneider, B., & Konz, A. M. (1989). Strategic job analysis. *Human Resource Management*, *28*, 51–63. doi:10.1002/hrm.3930280104

Schooler, C., Mulatu, M. S., & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging*, *14*, 483–506. doi:10.1037/0882-7974.14.3.483

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

Seibert, S. E., & Kraimer, M. L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior*, *58*, 1–21. doi:10.1006/jvbe.2000.1757

Shepherd, A. (2001). *Hierarchical task analysis*. New York, NY: Taylor & Francis.

Siddique, C. M. (2004). Job analysis: A strategic human resource management practice. *International Journal of Human Resource Management*, *15*, 219–244. doi:10.1080/0958519032000157438

Smith, G. F. (1997). Managerial problem solving: A problem-centered approach. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 371–380). Hillsdale, NJ: Erlbaum.

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*, 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*, 289–305. doi:10.1016/j.intell.2013.05.002

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. *Journal of Labor Economics*, *24*, 235–270.

Taylor, F. W. (1911). *The principles of scientific management*. New York, NY: Harper.

Taylor, W. G. K. (1989). The Kirton Adaption–Innovation Inventory: A re-examination of the factor structure. *Journal of Organizational Behavior*, *10*, 297–307. doi:10.1002/job.4030100402

Tushman, M., & O'Reilly, C. (2007). Research and relevance: Implications of Pasteur's quadrant for doctoral programs and faculty development. *Academy of Management Journal*, *50*, 769–774. doi:10.5465/AMJ.2007.26279169

Vancouver, J. B., & Schmitt, N. W. (2006). An exploratory examination of person-organization fit: Organizational goal congruence. *Personnel Psychology*, *44*, 333–352. doi:10.1111/j.1744-6570.1991.tb00962.x

Vargas Cortes, M. G., & Beruvides, M. G. (1996). An analysis of middle management work in non-steady conditions. *Computers & Industrial Engineering*, *31*, 53–57. doi:10.1016/0360-8352(96)00077-0

Vicente, K. J. (1999). *Cognitive work analysis: Toward safe, productive, and healthy computer-based work*. Hoboken, NJ: Taylor & Francis.

Warr, P., & Bunce, D. (1995). Trainee characteristics and the outcomes of open learning. *Personnel Psychology*, *48*, 347–375. doi:10.1111/j.1744-6570.1995.tb01761.x

Weede, E., & Kämpf, S. (2002). The impact of intelligence and institutional improvements on economic growth. *Kyklos*, *55*, 361–380. doi:10.1111/1467-6435.00191

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

Zaccaro, S. J., Mumford, M. D., Connelly, M. S., Marks, M. A., & Gilbert, J. A. (2000). Assessment of leader problem-solving capabilities. *The Leadership Quarterly*, *11*, 37–64. doi:10.1016/S1048-9843(99)00042-9

Zaccaro, S. J., Mumford, M. D., Marks, M. A., Connelly, M. S., Threlfall, K. V., Gilbert, J. A., & Fleishman, E. A. (1997). *Cognitive and temperament determinants of Army leadership* (U.S. Army Research Institute for the Behavioral and Social Sciences Tech. Rep. No. MRI 95-4). Bethesda, MD: Management Research Institute.

Zsambok, C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making*. Hillsdale, NJ: Erlbaum.