

Assessment and Theory in Complex Problem Solving ---A Continuing Contradiction?

Samuel Greiff

Department of Psychology, University of Heidelberg

Heidelberg University, Hauptstraße 47-51, 69117 Heidelberg, Germany

Tel: 49-622-154-7613 E-mail: samuel.greiff@psychologie.uni-heidelberg.de

Received: January 6, 2012

Accepted: March 14, 2012

Published: May 1, 2012

doi:10.5539/jedp.v2n1p49

URL: <http://dx.doi.org/10.5539/jedp.v2n1p49>

This research was supported by a grant of the German Research Foundation (DFG Fu 173/14-1).

Abstract

Complex Problem Solving (CPS) describes skills frequently needed in everyday life such as the use of new technological devices. Therefore, CPS skills constitute an increasingly important individual ability that needs theoretically embedded, reliable and validated measurement devices. The present article shows that current tests do not sufficiently address the requirement of a theory-based assessment. An integrative approach, the Action Theoretical Problem Space Model by Rollett (2008), is introduced and used to demonstrate how a theoretical framework can influence and inform test development. Implications for the assessment of CPS and its potential are discussed.

Keywords: Complex problem solving, Assessment, Action theories, Functionalist theories, Test development

1. Introduction

In the 1960ies, the majority of tasks at an average work place demanded almost exclusively routine cognitive skills. Today, 50 years later, routine tasks have almost entirely left the human workplace (Autor, Levy, & Murnane, 2003) and in the 21st century, humans all over the world are faced with increasingly complex tasks demanding flexibility and generalized problem solving skills. An obvious example of this is the rapidly rising prevalence of mobile phones. In fact, over 90% of adults in the Western world own a mobile phone and within the next years the number of registered mobile phones will exceed the global population. The advent of these devices has doubtlessly had a great impact on telecommunication and social interaction (Australian Research Council, 2007). In the 21st century, the ability to handle a mobile phone is taken for granted and today's generation seldom struggles with these kinds of requirements. On the other hand, people born 100 years ago would have been utterly lost confronted with a mobile phone. The process of becoming acquainted with it is characterized by non-routine and generalized problem solving skills. For the sake of illustration, imagine you came across a mobile phone for the first time ever. What would be your natural approach to master this device when blocking out prior knowledge completely?

First, you would press buttons (i.e., give inputs) in order to receive a reaction from the device (i.e., generate output; Greiff, 2012). From the observed connections between in- and outputs, you would acquire knowledge and generate a mental representation of the device (Markman, 1999). You could then use your knowledge to control it and to reach desired states (e.g., making a phone call; Novick & Bassok, 2005). These three aspects are referred to as *exploration*, *knowledge acquisition*, and *knowledge application* (Funke, 2001) and considered important aspects of complex problem solving (CPS; Dörner, 1986).

The objective of this paper is the concept of CPS, its theories and their neglected role in test development. After outlining a general understanding of CPS, I will revisit measurement approaches and some empirical findings leading me to yet unanswered questions. I will then consider theories of CPS and derive how CPS research may benefit from a theoretically motivated test development.

Generally, Mayer (1990) defines problem solving as the process of transforming a given state into a goal state when no means of solution is readily available. CPS is a specific form of problem solving and emphasizes a

problem solver's direct interaction with a previously unknown and dynamic system. As Buchner (1995) puts it, a problem is complex "if the problem situation changes as a function of user's intervention or as a function of time and the environment's regularities can only be revealed by successful exploration and integration of the information gained in that process" (p. 14). Due to the obvious importance of these skills in everyday life, large-scale assessments like the Programme for International Student Assessment (PISA) have realized the importance of CPS for educational contexts and included it in the 2012 assessment cycle. There, the need for an applied understanding is directly resembled in the definition: "Dynamic problem solving is the ability to identify the unknown structure of artifacts in dynamic [...] environments to reach certain goals [...]. In general, some exploration must be done to acquire the knowledge necessary to control the device" (OECD, 2010, p. 16). However, the varying terminology used for the construct under study has led to some confusion. Whereas the term *interactive problem solving* used in the PISA assessment reflects the inevitable interaction between problem solver and problem, Funke (2001) emphasized the dynamics inherent in each problem as the problem situation may change by itself over time (e.g., after a certain time without a user intervention the mobile phone may switch automatically into the stand-by mode) by introducing the term *dynamic problem solving*. However, the original term *complex problem solving* used throughout this paper was introduced by Dörner (1986) and refers to the complexity of the underlying system. That is, changing one variable in a task may lead to manifold changes in other variables. As this inconsistency in terminology was not necessitated by scientific reasons it has led to a considerable amount of confusion, which was further impeded by the width of problem solving research in general involving also problems in narrow and domain-specific areas such as mathematical, scientific, and technical problems (Sugrue, 1995) opposed to the domain-unspecific construct referred to as CPS. Whereas CPS may take place in technical contexts as shown in the mobile phone example, it is not limited to this type of context. Indeed, Funke (2001) and Novick and Bassok (2005) consider the processes underlying CPS as defining characteristics and not dependent on semantic or contextual embedding, which could be social, technical, or private (OECD, 2010). Please note that the mobile phone example only holds under the assumption that no prior knowledge is available because only then will the typical CPS processes of exploration, knowledge acquisition, and knowledge application be applied. Once sufficient knowledge is established, the mobile phone is mastered relying on prior experience. Funke (2003) denotes this a task but highlights that usually CPS is heavily involved when prior knowledge is gathered.

Even though the understanding of CPS is at its core domain-unspecific (cf. Buchner, 1995; Funke, 2001), some debate remains on whether there is a generic CPS ability or whether CPS largely depends on the specific domains involved and their content. Whereas some claim CPS to be domain-unspecific and thereby an overarching construct like general intelligence (Dörner, 1986; Funke & Frensch, 2007; Greiff, 2012), researchers in specific domains like math or science consider CPS to always be associated with specific contents and only slightly correlated between domains (cf. Sternberg & Frensch, 1991). Interestingly, research on experts and novices has also focused on differences within domains as diverse as reading, chess, or physics (e.g., Chi, Feltovich, & Glaser, 1981), but these domain-specific approaches have not been merged with general research on CPS. Sternberg (1995) was the first one to make this explicit: He criticized that research on complex problems has been focused too strongly on the comparison between experts and novices and specific differences between the two thereby overstating domain-specific processes and neglecting domain general processes that may be of considerable importance in education. Further, he criticized educational psychology a general neglect of CPS by stating that it "has not captured their [researchers and practitioners] imagination, at least not in the United States" (Sternberg, 1995, p. 300) and asked for comprehensive research on complex problems, which (1) are closer to real life (i.e., represent what happens in the classroom and is required in educational contexts) and (2) can be solved without specific expertise (i.e., also by students being considered novices). Indeed, these two points apply to the mobile phone example mentioned above and was the original idea when research on CPS was introduced in the 1970ies (cf. Frensch & Funke, 1995). In recent years, efforts to combine domain-specific (e.g., mathematical problem solving) and domain-unspecific (i.e., CPS) research on problem solving are still sparse and the issue remains largely unsolved.

In line with Sternberg's (1995) claims, there are two additional reasons that explain this: First of all, even though psychological and domain-specific research has existed for a long time, both lines largely ignore each other's findings. It seems that fundamentally different conceptualizations of problem solving either as a general and overarching cognitive ability or as a specific domain-bound ability of little generalizability do not trouble either research line and that the claims have been staked. The second reason is one within psychological research: CPS measurement devices lack a sufficient theoretical embedding and are largely constructed as *ad-hoc measures* thereby making it extremely difficult to go beyond the level of manifest variables when comparing the two lines of research. I refer to ad-hoc measures as measures that are constructed largely by face validity without a proper

definition of the construct under study, such as formal frameworks introduced by Funke (2001).

Funke (2001) used linear structural equation (LSE) models and finite state automata (FSA) to formally describe the underlying structure of different complex problems and to develop measures from the data produced while working on these problems. Complex problems based on the first formalism, LSE systems, are composed of quantitative connections between variables. More specifically, decreasing or increasing an input variable may in turn lead to a decrease or increase in one or several output variables (e.g., applying the volume up button several times on a mobile phone increases the volume accordingly). Problems based on the second formalism, FSA, on the other hand, are composed of qualitative connections between variables. More specifically, pressing one variable (e.g., a button) may transfer the system into a different state (e.g., pressing the OFF-button on the mobile phone changes the state of the entire device). Within each of these formalisms, a phase of exploration and knowledge acquisition is followed by a subsequent phase of knowledge application.

Both, LSE and FSA have been widely used to develop measures of CPS (e.g., Kluge, 2008; Kröner, Plass, & Leutner, 2005; Rollett, 2008). However, these formalisms do not constitute a theory taken by their own and were never intended to do so, but were supposed to offer a formal description of problem structures contributing to the development of measures. They have never been explicitly connected to a theoretical concept of CPS and neither have the measures deduced from them. The distinction of two phases, exploration and knowledge acquisition as the first and knowledge application as the second, which Funke (2001) suggested for reasons of simplicity, was not theoretically motivated. Nevertheless, many researchers took this delineation as given and as a replacement for a strong theoretical underpinning. Consequently, the construct captured within the formalisms remains blurry, which is also reflected in the amount of different tests published. Implicitly, it is assumed they all tap into the same construct, but considering their variety in complexity, underlying system structure, and the diversity of findings on the very same matters, I have doubts about that.

In fact, empirical findings on the comparability of different measures of CPS are scarce at best and a closer inspection of specific operationalizations leaves the impression of knowledge-lean approaches. In other words, the explicit or implicit definition of the construct is not sufficiently reflected in the measures. One could even argue that knowledge is gained only on the specific measures and not on CPS as a latent variable. This might - at least partly - explain why findings on the general or domain-specific nature of CPS have been exceptionally diverse in psychological research: For instance, Wittman and Süß (1999) explained CPS fully by prior knowledge and intelligence and concluded CPS was either non-existent or domain-specific. Further, correlations between different types of problem solving and domain-specific literacy, such as mathematics and reading in the PISA 2003 assessment also suggested the specific nature of CPS. On the other hand, Leutner, Klieme, Meyer, and Wirth (2004) reported strong dissociations between analytical problem solving and domain-specific literacy and Putz-Osterloh (1981) as well as Wüstenberg, Greiff, and Funke (2012) found empirical evidence for the general nature of CPS. This question of construct validity (i.e., is there a domain-unspecific CPS ability?) is directly connected to the question of how CPS relates convergently and divergently to other constructs such as intelligence and learning. Whereas the question whether intelligence and CPS are closely related, identical, or not related at all has been intensively disputed (cf. Rost, 2009), much less efforts have been made to distinguish learning and CPS, which may – after all – be of even higher importance. In fact, partly conflicting results on CPS may lie in a lack of a clear distinction between CPS and learning. Wirth, Küsting, and Leutner (2009) conceptually distinguish learning goals and complex problem solving goals. That is, learning goals are usually ill-defined internal mental states of a learner's knowledge, whereas goals in complex problems used in research are often well-defined and represented in the external environment. Thus, learning requires modifying internal schemata (i.e., structural knowledge), whereas CPS requires transforming some specific aspects of a problem solver's external environment in order to reach an externally represented goal state (i.e., instance knowledge). Wirth et al. (2009) and Küsting, Wirth, and Paas (2011) showed that this goal specificity effect (internal vs. external goals) affected learning and complex problem solving performance mediated by strategy and cognitive load. From this, one could conclude that presenting well-defined goals will help separating CPS from learning, but Kröner et al. (2005) argue that even when trying to reach a given and well-defined goal state problem solvers may yet learn important aspects of a problem's structure (i.e., structural knowledge), that is, they modify their internal schemata and not only their external environment. Further, within the formal frameworks of LSE and FSA introduced by Funke (2001) problem solvers are usually presented an unguided exploration phase called knowledge acquisition in which no specific goals are present but in which they are only instructed to explore the underlying system structure in line with the understanding of learning goals mentioned above. Does this imply that the first phase of knowledge acquisition within formal frameworks is heavily impacted by learning, whereas the second, knowledge application, in which specific goals are presented, is not? Even further, is this an example

how an atheoretical approach – as in LSE and FSA – unintentionally confuses different concepts such as learning and CPS? This conclusion would be premature, but it shows the confusion and confounding of concepts when it comes to a conceptually clear understanding of CPS, which is largely caused by the lack of theoretical embedment in measures of CPS. That is, from a theoretical perspective, learning goals should be excluded in the assessment of CPS by presenting only well-defined goals going along with Wirth et al. (2009).

Overall, the lack of theoretical embedment and the diversity of operationalizations are likely to contribute substantially to the empirically and conceptually confusing situation. But why have researchers abandoned theoretical considerations in the assessment of CPS? One likely reason is that not only psychological and domain-specific research have been kept apart but also within psychology, experimental and assessment research sometimes take little notice of each other. Whereas the former is mostly concerned with intergroup differences and theoretical considerations on CPS, the latter is emerging just now and addresses empirical aspects of reliability and validity. And yet, the missing link between assessment and theory is surprising as there are plenty of theories on CPS that could be applied under a measurement perspective. This is in line with Cronbach's claim (1957) that human cognition and learning can only be understood through the unification of experimental and differential approaches. I will now present theories on CPS eligible for an assessment application and provide ideas on how a theoretical approach may inform test development.

2. Theory-based Measurement in CPS

Already in the early 20th century Gestalt psychology and Psychoanalysis studied problem solving (for an overview see Funke, 2003), however, functionalist and action theories have dominated the research field in the last decades. From the functionalist perspective, mental states are characterized by their causal role in the process of problem solving, while action theoretical approaches state that actions generally are intentional and have to be seen in a broad context. Action theorists divide CPS in different phases and consider the entire process on a macro level, whereas functionalist theorists focus on the description, construction, and interaction of structural units on a micro level.

Both functionalist theories and action theories are fruitful approaches in describing complex problem solving processes (Fischer, Greiff, & Funke, 2012), so it seems logical to bring these two approaches together to overcome the specific weaknesses of the respective positions: The characteristic phases of CPS (which are described by action theories) could be augmented by process assumptions (which are proposed by functionalist approaches). A system acting meaningful according to certain reasons – the subject of action theories – may as well be described as being in a certain functional state that leads to the behavior/action observed (output) because of the environment perceived by the system (input) – that means as a subject of functionalism.

Especially for CPS, the phases proposed by action theories are of importance, as there are different demands for the problem solver in different phases of the process of CPS. That is, theoretically hypothesized phases generally need to be represented within a measure rendering it possible to empirically verify or falsify the underlying theoretical assumptions and to derive a meaningful assessment of CPS. On the other hand, one needs to go into much more detail than action theories currently do when measuring different processes of CPS, as an assessment does not only want to know what the characteristic phases of the course of CPS is, but how they are intertwined in detail and especially to what extent each phase is a necessary part of the CPS process. Therefore, one has to look at the processes involved in each of the phases and how they depend on each other.

No single approach seems to be fully adequate for this purpose: Weaknesses of the action theoretical approach are its generally vague level of description, its normative character, and its lack of considering structural units within the CPS process. In contrast to this, the functionalist approach focuses on the description, construction, and interaction of these structural units, but does not address the characteristic phases of CPS on a broader level and how the problem space changes over the course of problem solving. Thus, it seems logical to bring the two concepts together in order to integrate them within a coherent theoretical framework.

Rollett (2008) is the only one who so far has explicitly tried to integrate the functionalist and action theoretical approach. In his *Action Theoretical Problem Space Model (ATPSM)* he combines the functionalist three-space model of Vollmeyer and Burns (1999) and the action theoretical model of Schaub and Reimann (1999), even though this attempt has been up to now neither widely perceived nor taken up by test developers. In their functionalist model, Vollmeyer and Burns (1999) assume an *instant space* containing concrete states of a problem (e.g., turned-off mobile phone) and a *rule space* containing combinations of different states and the operations necessary to switch between states (e.g., pressing the green button turns the phone on). In an additional *model space*, assumptions about the rules (e.g., at no point two buttons have to be pressed simultaneously) are stored. Rollett (2008) describes a theory in which activity in these spaces and the interaction between them differs

depending on the current phase of the CPS process according to the six phases of Schaub and Reimann (1999): (1) *Goal elaboration* is associated with constructing a task model in the model space to reach a given goal. However, if a specific goal is not at hand, it is developed by an interaction between the instance space (concrete goals are elaborated) and the rule space (hypotheses on goal elaboration are tested). During the CPS process, the problem solver frequently refers back to this first stage. That is, Rollett (2008) does not consider CPS as serial process, but assumes that problem solvers frequently and unsystematically switch between phases. (2) *Modeling the system structure* is (a) based on a viable model of the problem, and (b) associated with an interaction of rule space (containing the structure between variables) and instance space (containing instances of the problem that are the basis for inferring its structure). During (3) *background control*, the representation of the task is simplified by excluding irrelevant information in instance and rule space. Phase (4), the planning and execution of *actions* is characterized by activity in the instance space in which states of the problem are altered by using suitable operators. These operators are selected in accordance with the rules stored in the rule space. During (5) *control* the difference between current state and goal state is tested in the instance space and the newly acquired findings lead to adjustments in the model and the rule space. In the final phase of (6) *updating*, consequences of the actions are considered and the representation of states, operators, and variables is updated in each space. The phases are repeatedly gone through until either the desired goal state is reached or the entire CPS process is aborted (Funke, 2011).

For the purpose of assessment, the six different CPS phases in the ATPSM need to be represented in specific operationalizations. This poses the additional challenge of translating theoretically proposed components into empirical scales, which in the end allow for an objective and reliable measurement of CPS, a challenge well known to psychological research. For instance, the rather obvious dissociation between theoretical definitions of intelligence and their according empirical operationalizations has led Boring (1923) to his seminal article *Intelligence as the test tests it* denouncing the common theoretical ground of intelligence tests. The common understanding of what CPS should be composed of is as diverse and varies as widely as for intelligence (Dörner, 1986) and may well have led to the aforementioned neglect of theory as it seemed impossible to adequately represent theoretical concepts of CPS in a test and to mirror them on a reliable scale. However, wouldn't the best way then be to accept that the theoretical understanding of CPS is narrowed down in specific operationalizations instead of largely obscuring theoretical considerations? Even further, I agree that CPS in its full width will be difficult to translate into empirical scales, but I disagree that this is not possible at all. Clearly, some (potentially narrow) aspects could be represented in a scalable test, but a theory is needed to do so. I will now argue which restrictions need to be accepted and how exemplary operationalizations of the six phases within the ATPSM may look like.

First of all, the phases assumed in the ATPSM have to be separated artificially, which poses a necessary restriction in order to make them accessible to measurement, but – at the same time – constrains the natural process of complex problem solving one would usually exhibit involving frequent switches between phases. The number and the nature of phases are driven by the underlying theory - in this example the ATPSM. Further, confounding variables such as prior knowledge need to be excluded to derive pure measures of CPS, for instance by choosing fictitious cover stories. In the real world, on the other hand, complex problem solving always involves prior knowledge as we hardly ever experience situations in which we have not at least some prior experience we can rely on and also in assessment one has to accept that a complete exclusion of prior knowledge is hardly possible. That is, even in the mobile phone example, a problem solver needs to know that the device can be used to make phone calls or to send text messages. However, when assessing CPS it is important that no specific content knowledge such as some laws of physics or particular mathematical operations are needed to solve the problem (OECD, 2010).

Given these restrictions posed on a full and externally valid representation of CPS in a test, the theoretical understanding within the ATPSM can still be translated into items. More specifically, in (1) *goal elaboration*, predefined goals will lead to a task model (in the model space) and specific goals (in the instance space) that could be assessed via open-ended questions. If goals are unspecific at the outset of the problem, they are clarified during the CPS process and can be assessed at different stages. A combination of both is achieved by an unspecific exploration phase and a control phase with predefined goals. Thus, the division between exploration and control inherent in some formal frameworks (Funke, 2001) receives a theoretical justification in the ATPSM. (2) *Modeling* could be readily measured through the knowledge of specific states of the problem in the instance space (e.g., “Is this a possible state of the problem?”; Kluge, 2008) and on the overall structure in the rule space (e.g., by causal diagrams, in which problem solvers have to give their representation of the underlying problem structure; Funke, 2001). How information is reduced during (3) *background control* is closely connected to Dörner's (1986) features of a problem situation. It can be assessed by either asking which inputs and outputs may be disregarded in a certain

task (e.g., the number keys on the mobile phone may be ignored when redialing) or by analyzing exploration patterns that may increasingly focus on certain variables and leave out others. Wirth (2004) describes this process as *identification* and *integration* and derives measures that identify more and less successful problem solvers. The operators chosen by the problem solver when applying (4) *actions* could be assessed through multiple-choice questions or in open-ended format. Additionally, testees could be asked why they preferred one operator to another adding a qualitative notion to the assessment. (5) *Control* and associated processes in the instance space may be measured by extrapolation and interpolation tasks (Wirth & Klieme, 2003) like “What happens if this button is pressed” or “Will applying this measure approach the desired goal state?” At the end of the CPS process, (6) *updating*, the final problem representation in rule and model space can be assessed formatively (Funke, 2001). The classification of rules, mental representation, and hypotheses on the task as right or wrong is comparable to status assessment applied in classical intelligence testing and yields comprehensive results on individuals’ performance.

3. Outlook and Implications

The present article has pointed towards the lack of theoretical considerations in the assessment of CPS and has simultaneously shown that a theoretical perspective is feasible. In fact, assessment concepts already in use such as measures used by Wirth and Klieme (2003) can be adapted to the ATPSM as shown above and even the practical distinction between a knowledge acquisition and a knowledge application phase inherent in Funke’s (2001) formalisms could be theoretically justified within the ATPSM. Indeed, one way to operationalize phase (2) *modeling the system structure* and phase (5) *control* in the ATPSM could utilize the distinction of knowledge acquisition and knowledge application as conceptualized in the LSE and FSA framework. Obviously, the ATPSM is one choice out of many and other theories could also provide sound starting points for a theoretically motivated test development. However, it may be an adequate choice as it provides a process-oriented understanding of CPS by incorporating features of both functionalist and action theoretical approach.

The question why researchers have largely ignored theoretical issues in the assessment of CPS remains unanswered. I can only speculate about this matter, but even Dörner (1986) has called for a theoretically embedded test and provided an associated theory and so has Rollett (2008). Nevertheless, both went only halfway failing to develop a corresponding test. Again, about the reasons I can only speculate. However, no speculation is needed to state that if measurement issues are sufficiently addressed, empirical results will inform the underlying theory, just as theory has informed test development, gaining important knowledge on the construct itself. For example, the empirical dimensionality of a CPS measure based on the six ATPSM phases may or may not support the assumption of these phases. That is, empirically a model composed of only three phases could be best to model the data (Greiff, 2012). The empirically obtained results, which are possibly further strengthened by cognitive modeling (Anderson & Lebiere, 1998), need to be related back to the theoretical understanding. Even further, only then can research questions on the nature of CPS be adequately addressed: Currently, the ATPSM understands CPS as an essentially domain-unspecific construct, but this would imply that performance in domain-specific problems is predicted not only by prior knowledge and measures of classical intelligence but incrementally by measures based on the ATPSM – a question yet open to empirical scrutiny.

The initial example of a mobile phone encountered for the first time and the individual processes associated with mastering it demonstrate how CPS is present in everyday life of the 21st century. In fact, we frequently encounter opaque systems and exploring, understanding, and controlling such a system is a prerequisite for successfully participating in today’s society. The ATPSM gives a theoretical understanding of the processes needed for effective complex problem solving. The real-world resemblance of CPS is in line with the intention of international large-scale assessments such as PISA: They are targeted at measuring abilities relevant in preparing students for society and a successful professional career rather than at measuring specific knowledge and rote learning in educational curricula. This explains the current interest of large-scale assessments in examining the set of skills involved in CPS as most important cognitive ability beyond intelligence (OECD, 2010). Interestingly, the understanding of CPS in PISA 2012 is already cross-curricular (i.e., domain-unspecific), whereas reading, mathematics, and science are inherently conceptualized as domain-specific problem solving. The understanding of cross-curricular problem solving will be widened in the PISA 2015 survey, in which collaborative aspects (e.g., communication and teamwork) within problem solving will be assessed posing exciting additional questions such as the role of peer-learning and parental teaching. Thus, PISA 2012 and 2015 will provide empirical data on domain-specific problem solving going beyond curricula in the content domains and, at the same time, on general and collaborative problem solving as cross-curricular abilities. Clearly, this conceptualization will confront the PISA survey with further cultural issues such as non-contextual effects of item familiarity, which may differ substantially across cultures. And yet or even more so, it is due time that psychological research on

complex problem solving substantiates this cross-curricular understanding of CPS by providing theoretically sound measures of CPS.

Pretz and Sternberg (2005) put out a passionate call for unified theories in cognitive sciences. In the field of CPS, however, one has to eat humble pie: Any theoretical perspective on assessment issues at all will be an advance to the status quo but I am optimistic that reconciling supposedly disparate perspectives will advance theory development in CPS, on one hand, and practice of assessment and interventions to promote skill acquisition in complex and dynamic domains, on the other. This, in turn, will lead to a better understanding of the role problem solving skills play in the beginning 21st century and the shift in demands associated with it.

References

- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. London: Lawrence Erlbaum Associates.
- Australian Research Council (2007). *The impact of mobile phones on work/life balance*. Unpublished report. Canberra, Australia: ARC.
- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118 (4), 1279-1333. <http://dx.doi.org/10.1162/003355303322552801>
- Boring, E. G. (1923). Intelligence as the tests test it. *New Republic*, 36, 35-37.
- 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.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152. http://dx.doi.org/10.1207/s15516709cog0502_2
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Dörner, D. (1986). Diagnostik der operativen Intelligenz [Assessment of operative intelligence]. *Diagnostica*, 32(4), 290-308.
- Fischer, A., Greiff, S., & Funke, J. (2012). The Process of Solving Complex Problems. *Journal of Problem Solving*, 4 (1), 19-42.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7, 69-89. <http://dx.doi.org/10.1080/13546780042000046>
- Funke, J. (2003). *Problemlösendes Denken* [Problem Solving Thinking]. Stuttgart: Kohlhammer.
- Funke, J. (2011). Problemlösen [Problem Solving]. In: T. Betsch, J. Funke, & H. Plessner (Eds.). *Denken – Urteilen Entscheiden Problemlösen* (p. 135-199). Heidelberg: Springer.
- Funke, J., & Frensch, P. A. (2007). CPS: The European perspective - 10 years after. In D. H. Jonassen (Ed.), *Learning to solve complex scientific problems* (pp. 25-47). New York: Lawrence Erlbaum.
- Greiff, S. (2012). *Individualdiagnostik der Problemlösefähigkeit* [Individual diagnostics of problem solving competency]. Münster: Waxmann.
- Kluge, A. (2008). Performance assessment with microworlds and their difficulty. *Applied Psychological Measurement*, 32 (2), 156-180. <http://dx.doi.org/10.1177/0146621607300015>
- Kröner, S., Plass, J. L., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, 33 (4), 347-368.
- Künsting, J., Wirth, J., & Paas, F. (2011). The goal specificity effect on strategy use and instructional efficiency during computer-based scientific discovery learning. *Computers & Education*, 56, 668-679.
- Leutner, D., Klieme, E., Meyer, K., & Wirth, J. (2004). Problemlösen [Problem solving]. In PISA-Consortium Germany, *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (pp. 147- 175). Münster: Waxmann.
- Markman, A. B. (1999). *Knowledge representation*. Mahwah, NJ: Erlbaum.
- Mayer, R. E. (1990). Problem solving. In M. W. Eysenck (Ed.), *The Blackwell Dictionary of Cognitive Psychology* (pp. 284–288). Oxford: Basil Blackwell.
- Novick, L. R., & Bassok, M. (2005). Problem solving. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (p. 321-349). Cambridge, NY: University Press.

OECD (2010). *PISA 2012 problem solving framework*. Paris: OECD.

Pretz, J. E., & Sternberg, R. J. (2005). Unifying the field. Cognition and intelligence. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition & Intelligence. Identifying the mechanisms of the mind* (pp. 306-318). Cambridge: University Press.

Putz-Osterloh, W. (1981). Über die Beziehung zwischen Testintelligenz und Problemlöseerfolg [On the relationship of test intelligence and problem solving success]. *Zeitschrift für Psychologie*, 189, 79-100.

Rollett, W. (2008). Strategieeinsatz, erzeugte Information und Informationsnutzung bei der Exploration und Steuerung komplexer dynamischer Systeme [Strategy usage, generated information and used information in the exploration and control of complex dynamic systems]. Münster: LIT-Verlag.

Rost, D. H. (2009). *Intelligenz* [Intelligence]. Weinheim: PVU Beltz.

Schaub, H., & Reimann, R. (1999). Zur Rolle des Wissens beim komplexen Problemlösen [On the influence of knowledge to complex problem solving]. In H. Gruber, W. Mack, & A. Ziegler (Eds.), *Wissen und Denken. Beiträge aus Problemlösepsychologie und Wissenspsychologie* (pp. 169-191). Wiesbaden: Deutscher Universitäts Verlag.

Sternberg, R. J. (1995). Expertise in complex problem solving: A comparison of alternative concepts. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving. The European perspective* (pp. 295-321). Hillsdale, NJ: Lawrence Erlbaum.

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

Sugrue, B. (1995). A theory-based framework for assessing domain-specific problem-solving ability. *Educational Measurement Issues and Practice*, 14 (3), 29-36.

Vollmeyer, R., & Burns, B. D. (1999). Problemlösen und Hypothesentesten [Problem solving and hypothesis testing]. In H. Gruber, W. Mack & A. Ziegler (Eds.), *Wissen und Denken. Beiträge aus Problemlösepsychologie und Wissenspsychologie* (pp. 101-118). Wiesbaden: Deutscher Universitäts Verlag.

Wirth, J. (2004). *Selbstregulation von Lernprozessen* [Self-regulation of learning processes]. Münster: Waxmann.

Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy, & Practice*, 10, 329-345. Wirth, J., Küsting, J., & Leutner, D. (2009). The impact of goal specificity and goal type on learning outcome and cognitive load. *Computers in Human Behavior*, 25, 299-305.

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* (pp. 77-108). Washington: American Psychological Association.

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