



Assessing analytic and interactive aspects of problem solving competency



Andreas Fischer^{a,*}, Samuel Greiff^b, Sascha Wüstenberg^b, Jens Fleischer^c, Florian Buchwald^c, Joachim Funke^a

^a University of Heidelberg, Hauptstr. 47-50, 69117 Heidelberg, Germany

^b University of Luxembourg, 6, rue Richard Coudenhove Kalergi, 1359 Luxembourg-Kirchberg, Luxembourg

^c University of Duisburg-Essen, Berliner Platz 6-8, 45127 Essen, Germany

ARTICLE INFO

Article history:

Received 9 May 2014

Received in revised form 29 January 2015

Accepted 20 February 2015

Keywords:

Interactive Problem Solving

Analytic Problem Solving

Complex Problem Solving

Reasoning

Assessment

MicroDYN

ABSTRACT

This study is about different approaches to assessing Problem Solving Competency (PSC) applied in international large-scale assessments: Analytic Problem Solving (APS) and Interactive Problem Solving (IPS). Based on a university student sample ($n = 339$) and a high-school student sample ($n = 577$) we found that both approaches are highly interrelated in both samples, even after controlling for reasoning ($R^2 = .33$ to $.52$) indicating that both approaches address a common core of PSC. However, our results also indicate that unique aspects of APS and IPS (beyond each other and reasoning) are explanatory for school achievements in the high-school student sample. However, in the university student sample, only APS has a unique contribution to explaining school achievements (beyond IPS and reasoning) and our findings indicate, that APS – and not interactivity itself – may explain the incremental validity of IPS (beyond reasoning) reported in previous studies. Implications for problem solving research and educational practice are discussed.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Solving real problems is a complex endeavor: Even the most intelligent persons can fail solving realistic and complex problems, if they don't have important content knowledge or don't know adequate search strategies as well as when to apply them in an adaptive way (cf. Dörner, 1996; Fischer, Greiff, & Funke, under review). This paper is about some of the most important components of Problem Solving Competency (PSC, cf. Fleischer, Wirth, & Leutner, 2014; Greiff & Fischer, 2013a; Wirth & Klieme, 2003) and their interrelations. Problem Solving Competency can be understood as the ability to figure out a solution method for reaching ones goal if no such method is obvious (cf., Duncker, 1945; Wirth & Klieme, 2003), that is, to represent and solve problems in various domains (cf. Bassok & Novick, 2012; Schoppek &

Putz-Osterloh, 2003). In international large-scale assessments two different kinds of problems have been proposed for assessing PSC (OECD, 2014):

- 1) One kind of problem requires a single choice of a solution based on the information given at the outset. A characteristic example for this kind of problem is the problem of finding the shortest path between a set of locations based on a map before actually starting to travel. Problems of this kind can be solved analytically, as all the information required for finding a solution is given at the outset of the problem. We will refer to this kind of problem solving as *Analytic Problem Solving (APS)*.
- 2) The other kind of problem requires a series of multiple choices, where later choices can be influenced by the results of previous choices (also known as Dynamic Decision Making, e.g., Gonzalez, Lerch, & Lebiere, 2003). For instance, after starting a travel, the initial plan of which locations to see may be adapted dynamically to unforeseen changes in the situation (e.g., road works on certain paths). In this kind of problem, the problem solver can adapt his or her initial plans and knowledge at multiple points in time, because there is feedback after each interaction with the problem. We will refer to this kind of problem solving as *Interactive Problem Solving (IPS)*.

* Corresponding author at: Hauptstr, 47-50, 69117 Heidelberg, Germany. Tel.: +49 6221 547301; fax: +49 6221 547273.

E-mail addresses: andreas.fischer@psychologie.uni-heidelberg.de (A. Fischer), samuel.greiff@uni.lu (S. Greiff), sascha.wuestenberg@uni.lu (S. Wüstenberg), jens.fleischer@uni-due.de (J. Fleischer), florian.buchwald@uni-due.de (F. Buchwald), joachim.funke@psychologie.uni-heidelberg.de (J. Funke).

Both kinds of problems¹ have been proposed to measure PSC, but up to now it has never been tested conclusively, if performance in both measures (APS and IPS) indicates distinct facets of PSC, or if they can be considered to address a common core of PSC (e.g., strategies for analyzing complex problem statements, or for systematically structuring prior knowledge and complex information in a goal-oriented way) sufficiently distinct from logical reasoning (Raven, 2000). In the *Programme for International Student Assessment (PISA) 2012*, both kinds of problems have been used to assess a single underlying PSC factor (OECD, 2014). The studies of Wirth and Klieme (2003) and Scherer and Tiemann (2014) presented first evidence for a multidimensional structure of PSC but they did neither control for reasoning nor analyze external validity of the facets reported.

In the current paper we will clarify the conceptual interrelations of reasoning and PSC and we will present empirical evidence based on two samples (577 high-school students and 339 university students) to demonstrate that APS and IPS address a common core of PSC that cannot be explained by reasoning, and that APS and IPS additionally address unique aspects each, which are important for explaining external criteria beyond reasoning. In the discussion we will focus on findings consistent between samples.

1.1. (Why) PSC is conceptually different from reasoning

It seems obvious that basic logical reasoning (e.g., forming inductive or deductive conclusions based on facts or premises, cf. Carpenter, Just, & Shell, 1990; Mayer, 2011), is closely related to problem solving (Mayer, 2011) and necessarily involved in each valid approach to assess PSC (cf. Greiff & Fischer, 2013a; Wüstenberg et al., 2012). However, in addition to this kind of reasoning PSC also implies a large amount of crystallized² abilities (Postlethwaite, 2011), that is, “the knowledge and language of the dominant culture” (Horn & Masunaga, 2006, p. 590). More specifically, solving problems in a competent way involves “experimental interactions with the environment” (Raven, 2000, p. 54) and depends on a large base of procedural and declarative knowledge on *how* and *when* to perform different search strategies in order to adequately represent and solve problems (e.g., Dörner, 1996). The importance of crystallized knowledge, especially knowledge about strategies, for PSC has often been emphasized (e.g., Scherer & Tiemann, 2014; Schoppek & Putz-Osterloh, 2003; Strohschneider & Guss, 1999; Tricot & Sweller, 2014) and is a central conceptual difference to basic logical reasoning.³

If this claim is correct, each valid operationalization of PSC should prove to be *incrementally* valid, compared to tests of reasoning with regard to external criteria such as academic or occupational success. To our knowledge, it is an open question if common variance between current instances of Analytic and Interactive Problem Solving (e.g., Scherer & Tiemann, 2014) can be attributed to reasoning only.

The present study aims to clarify if both APS and IPS are valid approaches to assessing PSC, that is, if they address “more than reasoning” (Wüstenberg et al., 2012) with regard to explaining (1) each other or (2) school grades (as external criteria of PSC).

1.2. Concept and empirical results concerning Analytic Problem Solving

For a long time, PSC has been assessed by APS tasks, that is, by confronting participants with multiple heterogeneous problems each requiring a single solution to be generated analytically (e.g., Boggiano, Flink, Shields, Seelbach, & Barrett, 1993; Fleischer, Buchwald, Wirth, Rumann, & Leutner, under review; Fleischer, Wirth, Rumann, & Leutner, 2010; OECD, 2003). For instance, in PISA 2003 PSC was assessed by a set of multiple problems (OECD, 2003) that required (1) decision making under constraints, (2) evaluating and designing systems for a particular situation, or (3) trouble-shooting a malfunctioning device or system based on a set of symptoms (OECD, 2004, p. 61). All problems were designed to be realistic and refer to “cross-disciplinary situations where the solution path is not immediately obvious and where the literacy domains or curricular areas that might be applicable are not within a single domain of mathematics, science or reading” (OECD, 2003, p. 156; see also Leutner, Funke, Klieme, & Wirth, 2005a,b; Leutner, Wirth, Klieme, & Funke, 2005b).

Empirically, APS is highly correlated to performance in different domains like mathematics ($r = .89$), reading ($r = .82$) and science ($r = .80$) on a latent level (OECD, 2004, p. 55). Due to its broad operationalization APS is also closely related to – but yet empirically distinct from – reasoning ($r = .72$; Leutner, Klieme, Meyer, & Wirth, 2004; $r = .67$, Leutner, Fleischer, & Wirth, 2006; $r = .60$ Scherer & Tiemann, 2014). In general, APS seems to be more strongly related to intelligence and school achievements than IPS is (cf. Leutner et al., 2005a,b; Leutner, Fleischer, Wirth, Greiff, & Funke, 2012; Wirth & Klieme, 2003). To our knowledge there is no study explicitly examining the incremental value of APS over and above measures of reasoning and IPS.

1.3. Concept and empirical results concerning Interactive Problem Solving

IPS tasks are a more recent and computer-based approach to assessing PSC that evolved from research on Complex Problem Solving and Dynamic Decision Making (cf. Fischer, Greiff, & Funke, 2012). The defining feature of IPS is that the problem solver can not only rely on the information given at the outset, but must adapt his or her hypotheses (about how the problem works) and plans (about how to reach one's goals) while interacting with the problem (cf. Fischer et al., 2012; Klahr, 2000). Thus, the IPS approach focuses on effective strategies for searching the spaces of information and hypotheses as well as the resulting problem space (Greiff et al., 2013b). Fig. 1 illustrates an example of a typical interactive problem: This problem is an interactive computer-simulation based on a complex⁴ abstract linear equation model (cf. MicroDYN approach, Greiff, 2012; Greiff, Fischer, Stadler, & Wüstenberg, in press). It is about a handball-team, that can be trained by applying different amounts of three different trainings (labeled A, B, & C), with each training possibly influencing motivation, power of throw and exhaustion of the team. The problem has to be solved in two subsequent phases: In a first phase, the problem solver can vary the values of certain input variables (in this case representing the amounts of three trainings, shown on the left side of the screen in Fig. 1), and observe the values of certain output variables (on the right side of the screen in Fig. 1). In this phase, his or her goal is to find out about the causal structure of the simulation and to draw his or her hypotheses into a causal model at the bottom of the screen (*problem representation*, sometimes referred to as knowledge acquisition, see Fig. 1). In a subsequent phase the problem solver is instructed to reach a set of well-defined goals (see the values in brackets in Fig. 1) by

¹ In the literature on complex problem solving (e.g., Funke, 2003; Scherer & Tiemann, 2014) and dynamic decision making (e.g., Edwards, 1962), sometimes APS and IPS have also been referred to as static vs. dynamic decision problems, or as simple vs. complex problems, respectively.

² Traditional measures of “crystallized intelligence” are often tests of highly *general declarative* knowledge. They focus on breadth instead of depth of the individual's knowledge base (i.e., they “measure only the elementary knowledge, the beginning [declarative] knowledge, in the various fields of human culture”, Horn & Masunaga, 2006, p. 597). The *concept* of crystallized intelligence represents a broader and more diverse range of knowledge (Horn & Masunaga, 2006) – e.g., procedural knowledge as it is tapped by some tests of expertise or PSC, for example.

³ As a result of these crystallized aspects, PSC can be assumed to be less domain-general than reasoning as well as more prone to training (cf. Scherer & Tiemann, 2014).

⁴ Of course one could also simulate even more complex problems containing aspects like negative feedback (e.g., predator-prey-systems, Cushing, 1977; or the sugar-factory-simulation, Berry & Broadbent, 1984), phase transitions, or deterministic chaos (e.g., Verhulst, 1839) within the framework proposed by Funke (2001) but each of these aspects again is likely to address additional or different skills and strategies. Traditional MicroDYN tests seem to reliably address a small set of skills (cf. Greiff & Fischer, 2013a,b; Funke, 2010), that are central for solving a wide range of analytic and/or complex problems.



Fig. 1. Screenshot of a MicroDYN simulation. Input-variables (left side of the screen) are labeled with fictional pseudo-words, in order to not trigger any helpful prior knowledge about the problem's causal structure. In a first phase, a problem solver has to interact with the problem and draw the causal structure in a causal diagram (representation at lower part of the screen). Afterwards, in a second phase, certain goal ranges are shown for each output-variable (solution at the right side of the screen).

specifying a series of inputs (*problem solution* sometimes referred to as knowledge application, see Fig. 1).

Recent empirical studies shed light on the aspects of PSC that are assessed within this operationalization of the IPS approach: In the first phase, finding an adequate problem representation seems to primarily depend on applying the control-of-variables strategy, that is, varying one thing at a time ($r = .97$ on a latent level; Wüstenberg et al., 2012), and it seems to indicate a thoughtful application of adequate strategies in the dual-search of hypotheses and information (Greiff et al., 2013b). In the second phase of IPS, finding a solution primarily depends on the application of basal strategies for searching well-defined problem spaces, that is, functional equivalents of means–end analysis (Greiff et al., 2013b; Simon, 1975). The strategies for solution in IPS are highly similar to the ones involved in solving tests of reasoning. Correspondingly, most studies on the incremental validity of the IPS approach demonstrated incremental validity over and above different measures of reasoning *only* for the representation but not for solution in IPS (Greiff & Fischer, 2013a; Greiff et al., 2013b; Wüstenberg et al., 2012).

1.4. Hypotheses

In the current paper, we will test hypotheses regarding two main research questions: (1) do APS and IPS address a common core of Problem Solving Competency that cannot be explained by reasoning? And (2) do APS and IPS additionally address unique aspects of external criteria each?

With regard to the first research question we expect APS to be correlated to both facets of IPS due to a common impact of reasoning (Hypothesis 1a). Furthermore, when regressing APS on both facets of

IPS as well as on reasoning, we expect a unique contribution of problem representation in IPS (Hypothesis 1b) due to the additional impact of PSC on both IPS and APS, but we expect no unique contribution of problem solution in IPS because the search for a solution is very closely related to reasoning (Hypothesis 1c).

With regard to the second research question, we will examine the external validity of the IPS and the APS approach: We expect both approaches to be predictive for school grades as external criteria of Problem Solving Competency (Hypothesis 2a). More specifically, when regressing school grades on APS and IPS as well as on reasoning we expect a unique contribution of problem representation in IPS (Hypothesis 2b) because the strategic knowledge indicated by the representation may be important for school grades even beyond APS due to interactive aspects (cf. Wüstenberg et al., 2012). Again, we expect no unique contribution of problem solution in IPS because of its close relation reasoning (Hypothesis 2c). Most of all we expect a unique contribution of APS (Hypothesis 2d), because APS is known to be more closely related to school grades in different domains (cf. Scherer & Tiemann, 2014).

2. Method

2.1. Participants

We assessed two samples: One sample consists of 577 German undergraduate high-school students (age: $M = 14.94$; $SD = 1.29$; gender: 44.2% male, 46.8% female, 9% missing), mainly in the 9th grade (45.1%) but also from 8th grade (26.7%), 10th grade (14.6%) and 11th grade (7.4%) – with 6% missing an entry. Participants in the high-school student sample received 50 Euro for their class inventory. Two samples were chosen to ensure a sufficient amount of generality concerning

our findings (factorial invariance of IPS holds across grade levels 5 and 12, e.g., Greiff et al., 2013c; Scherer & Tiemann, 2014).

The second sample consisted of 339 German university students (age: $M = 22.30$, $SD = 4.02$; gender: 27.1% male, 67.6% female; 5.3% missing) from different fields of study. Most participants of this sample studied social sciences (57%), some participants studied natural sciences (28%) or had a different field of study (15%). Participants in this sample received either 25 Euro for participation or 4 h course credit.

The relations of APS, IPS, reasoning and GPA are original to this paper.⁵

2.2. Materials

2.2.1. Interactive Problem Solving (representation and solution)

The IPS approach to measuring Problem Solving Competency was operationalized by the MicroDYN test of Problem Solving Competency, based on a set of 10 tasks in each sample. According to the detailed description in Section 1.3 of this paper each task consisted of searching for *representation* and *solution* as two subsequent interactive subtasks (for a detailed description of MicroDYN refer to Wüstenberg et al., 2012; Greiff & Fischer, 2013a,b).

For each task, both the correctness of the causal diagram after phase 1 (representation) and the value of output-variables after phase 2 (solution) were scored dichotomously as either (1) *correct* or (0) *incorrect* (cf. Wüstenberg et al., 2012) to indicate strategic knowledge for finding adequate representations and solutions to minimal-complex interactive problems.

2.2.2. Analytic Problem Solving

The APS approach to measuring Problem Solving Competency was based on a set of realistic static problems that were applied in PISA 2003. In the university student sample we applied 4 items (Transit System Q1, Holiday Q2, Course Design Q1 and Freezer Q2) whereas in the high-school student sample we applied 6 items (Cinema Outing Q1, Cinema Outing Q2, Irrigation System Q3, Holiday Q2, Transit System Q1, Childrens' Camp Q1). A detailed description of items can be found in OECD (2004, pp. 59 ff.). The approach is elaborated in more detail in Section 1.2.

For some problems, answers were scored dichotomously as *incorrect* (0) or *correct* (1), for some problems answers were scored polytomously as completely *incorrect* (0), *partially correct* (1), or *correct* (2).

2.2.3. Reasoning (subtest of I-S-T 2000 R; KFT 4-12 + R)

In the university student sample, reasoning was assessed using the matrix subtest of the "Intelligence Structure Test-Revised" (I-S-T 2000 R; Liepmann, 2007). This test consisted of 20×2 -matrices, each containing a figural stimulus in each but one cell. In each matrix, one stimulus was missing, and participants had to choose the missing figural stimulus out of five alternatives. Answers were scored as right or wrong. Missing values were considered wrong answers.

In the high-school student sample reasoning was assessed by the subscale "figural reasoning" of the KFT 4-12 + R (Heller & Perleth, 2000). This test consists of 23 items requiring to identify the relation of a pair of two figures and to choose one out of five alternatives in order to complete a second pair of figures with the same relation. Answers were scored as right or wrong. Missing values were considered wrong answers.

⁵ Other data of the university student sample was published in Greiff et al., (2013b), who studied complex problem solving as a latent construct, with the MicroDYN test being one of multiple indicators of complex problem solving. Part of the MicroDYN- and reasoning-data of a subsample of the high-school student data was published in Frischkorn, Greiff and Wüstenberg (2014) who studied the development of complex problem solving.

2.2.4. School grades

Subjects were asked to report their final Grade Point Average (GPA; university student sample) or their last grades in each course (high-school student sample). In the high-school student sample we used a latent factor model to estimate the current GPA. As grades in the German school system range from 1 (*very good*) to 6 (*insufficient*), we reversed GPA, so that higher numerical values indicate better performance.

2.3. Procedure

The university students were tested in two sessions (each lasting about 120 min). In the first session participants worked on the IPS test (about 60 min) and the APS test (about 15 min). In the second session they worked on a set of tests including the reasoning task (about 10 min).

The high-school students were tested in two sessions (45 min each). In the first session participants worked on the IPS test and in the second session on the APS test (about 35 min) and the reasoning tasks (about 10 min).

3. Results

All latent analyses were obtained by using MPlus 5.21 (Muthén & Muthén, 2008), commonality analysis was run in Gnu R by using the yhat package (Rya-Mukherjee et al., 2014). WLSMV-estimators were chosen for the structural equation models with ordinal items (cf. Muthén & Muthén, 2007).

3.1. Measurement models

This section specifies measurement models for each latent construct. With regard to fit indices Hu and Bentler (1999) recommend to use models with Comparative Fit Index (CFI) and a Tucker Lewis Index (TLI) value above .95 and a Root Mean Square Error of Approximation (RMSEA) below .06.

IPS was modeled as a construct with the two correlated factors representation and solution, which fitted the data well (CFI $\geq .955$, TLI $\geq .970$; RMSEA $\leq .053$). APS was modeled as a one-dimensional construct (CFI $\geq .949$; TLI $\geq .923$; RMSEA $\leq .054$) in accordance with the modeling procedure in PISA 2003 (OECD, 2005). Cronbach's Alpha for APS was comparatively low (see Tables 1 and 2) which is consistent with PISA 2003 (OECD, 2005; Fleischer et al., 2014) and with prior studies on APS (Fleischer et al., 2014). Reasoning was also modeled as a one-dimensional construct (CFI $\geq .96$; TLI $\geq .988$; RMSEA $\leq .046$) according to the test manual. In the university student sample each item of the reasoning test was assigned to one of four parcels as described by Greiff et al. (2013b) (CFI = 1.00; TLI = 1.00; RMSEA < .001). In the high-school student sample Grade Point Average was modeled – with slightly suboptimal fit – as a one-dimensional construct indicated by current grades in German, English, Math, Physics, Chemistry, and Biology (CFI = .946; TLI = .953; RMSEA = .166).

Table 1

Latent correlations of both aspects of IPS (representation and solution), APS, reasoning and grade point average (GPA) in the high-school student sample.

	Representation	Solution	APS	Reasoning	GPA
Representation	$\alpha = .807$				
Solution	.777***	$\alpha = .766$			
APS	.747***	.729***	$\alpha = .543$		
Reasoning	.520***	.469***	.562***	$\alpha = .914$	
GPA	.341***	.226***	.379***	.264***	$\alpha = .811$

Note. α : Cronbach's Alpha; $n = 577$.

*** $p < .001$.

Table 2
Latent correlations of both aspects of IPS (representation and solution), APS, reasoning and grade point average (GPA) in university student sample.

	Representation	Solution	APS	Reasoning	GPA
Representation	$\alpha = .792$				
Solution	.846***	$\alpha = .765$			
APS	.768***	.763***	$\alpha = .462$		
Reasoning	.214***	.273***	.395***	$\alpha = .820$	
GPA	.209***	.206***	.395***	.118*	–

Note. α : Cronbach's Alpha; $n = 339$.
* $p < .05$.
*** $p < .001$.

3.2. Structural models

In order to address our first research question (Section 3.2.1) we ran latent regressions of APS on representation, solution, and reasoning for the high-school student sample ($R^2 = .643$; CFI = .966; TLI = .981, RMSEA = .035) and for the university student sample ($R^2 = .672$; CFI = .963; TLI = .975, RMSEA = .038). In order to address our second research question (see Section 3.2.2) we ran latent regressions of GPA on representation, solution and reasoning for the high-school student sample ($R^2 = .172$; CFI = .964; TLI = .979, RMSEA = .035) and the university student sample ($R^2 = .186$; CFI = .963; TLI = .975, RMSEA = .037). Explained variance was highly significant in all models ($p < .01$) and variance-inflation was not indicated (variance-inflation factors below 5 for all predictors, cf. O'Brien, 2007).

Latent correlations between all constructs assessed (see Tables 1 and 2) proved to be positive and substantial for all measures.

Additionally we ran commonality analyses (Nimon, Lewis, Kane, & Haynes, 2008; Rya-Mukherjee et al., 2014) for each of these regression models in order to decompose the explained variance of the criterion into unique and shared contributions for each predictor (Rya-Mukherjee et al., 2014).

3.2.1. Research question 1: does APS share variance with the facets of IPS over and above reasoning?

Our first research question was whether the IPS facets representation and solution were predictive for APS performance over and above reasoning. The regression of APS on representation, solution, and reasoning (see Fig. 2 for the high-school sample) indicated significant unique contributions of representation ($\beta = .372$ –.438; $p < .05$) and reasoning ($\beta = .209$; $p < .01$) in both samples. The unique contribution of solution was significant in the high-school student sample ($\beta = .341$; $p < .01$) but not in the university student sample ($\beta = .334$; $p = .10$).

The corresponding commonality analyses (see Tables 3 and 4) revealed that a large amount of the explained variance in APS could be attributed to variance that is common to representation and solution (and not common to reasoning) in both samples ($R^2 = .23$ –.43). Additionally,

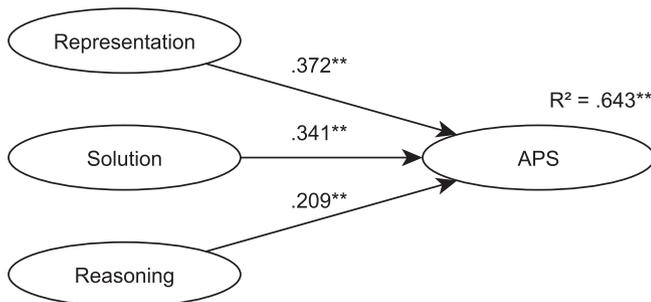


Fig. 2. Unique contributions of IPS representation, IPS solution and reasoning in predicting APS (model b) in the high-school student sample. For correlations between predictors see Table 2. **: $p < .01$.

Table 3
Unique and common commonality coefficients ΔR^2 and the corresponding percent of explained variance (%total) for each predictor in the regression of APS on representation, solution, and reasoning based on the high-school student sample.

Variables	ΔR^2	%total
Unique to representation	0.051	7.93
Unique to solution	0.045	7.04
Unique to reasoning	0.031	4.80
Common to representation, and solution	0.232	36.04
Common to representation, and reasoning	0.031	4.83
Common to solution, and reasoning	0.010	1.60
Common to representation, solution, and reasoning	0.243	37.76
Total	0.645	100.00

a substantial amount of explained variance in APS could be attributed to variance common to representation, solution and reasoning in both samples ($R^2 = .10$ –.24).

In summary, we found substantial commonalities between IPS, APS and reasoning (supporting Hypothesis 1a) as well as a unique contribution of IPS representation to explaining APS, beyond IPS solution and reasoning (supporting Hypothesis 1b). We did not find consistent evidence for a unique contribution of IPS solution (partially supporting Hypothesis 1c).

3.2.2. Research question 2: can GPA be explained by unique aspects of APS or the facets of IPS over and above reasoning?

Our second research question was whether there were unique contributions of APS and the facets of IPS over and above reasoning to predicting school grades as an external criterion of Problem Solving Competency. The regression of GPA on representation, solution, APS, and reasoning (see Fig. 3 for the high-school sample) indicated a significant unique contribution of APS ($\beta = .336$ –.644; $p < .05$) and no significant unique contribution of reasoning ($\beta = -.067$ –.060; $p = .392$ –.352) in both samples. The unique contributions of representation and solution were significant in the high-school student sample ($\beta = .241$ to $-.234$; $p < .05$) but not in the university student sample ($\beta = -.159$ to $-.133$; $p = .43$ to $.51$).

The corresponding commonality analyses (see Tables 5 and 6) underlined the importance of APS as they revealed a substantial amount of the explained variance in GPA ($R^2_{Total} = .17$ –.19) could be attributed to variance that is unique to APS in both samples ($R^2_{APS} = .04$ –.14). Please note, negative coefficients in Table 5 are not problematic for the analysis: Given the positive correlations among all predictors they indicate statistical suppression effects related to solution in the high-school sample (Rya-Mukherjee et al., 2014).

In summary, we found substantial correlations between GPA and both IPS and APS (supporting Hypothesis 2a) as well as a unique contribution of APS to explaining GPA, beyond IPS and reasoning (supporting Hypothesis 2d). In this regard, we did not find consistent evidence for a unique contribution of IPS representation or IPS solution (partially supporting Hypotheses 2b and 2c).

Table 4
Unique and common commonality coefficients ΔR^2 and the corresponding percent of explained variance (%total) for each predictor in the regressions of APS on representation, solution, and reasoning based on the university student sample.

Variables	ΔR^2	%total
Unique to representation	0.056	8.26
Unique to solution	0.030	4.45
Unique to reasoning	0.041	6.03
Common to representation, and solution	0.434	64.19
Common to representation, and reasoning	–0.003	–0.45
Common to solution, and reasoning	0.015	2.23
Common to representation, solution, and reasoning	0.103	15.29
Total	0.676	100.00

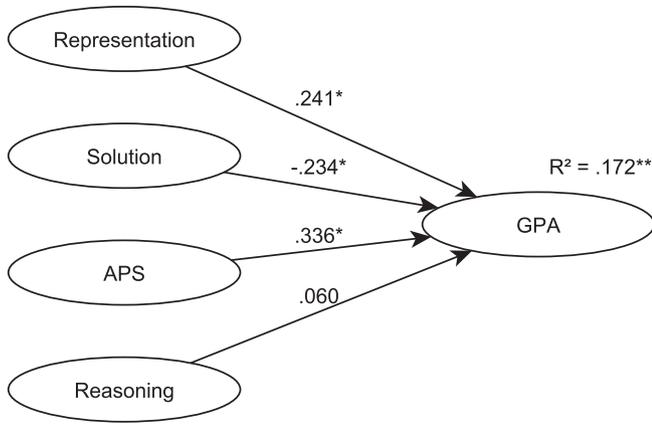


Fig. 3. Unique contributions of IPS representation, IPS solution, APS and reasoning to predicting GPA (model c) in the high-school student sample. For correlations between predictors see Table 1. *: $p < .05$; **: $p < .01$.

4. Discussion

In the current paper we have outlined and contrasted two different approaches to assessing aspects of Problem Solving Competency: One approach is based on static problems that have to be solved analytically (i.e., APS), whereas the other approach is about dynamic decisions in problem situations which have to be represented and solved interactively (i.e., IPS).

4.1. Research question 1: does APS share variance with the facets of IPS over and above reasoning?

Our first question was whether performance in APS shares variance with the crystallized strategic knowledge assessed by IPS. Indeed, we found APS highly correlated to both facets of IPS, and commonality analysis revealed that – besides a medium to large amount of variance that APS shares with IPS and reasoning (supporting Hypothesis 1a) – a large amount of variance in APS can be explained by the facets of IPS (over and above reasoning). Especially representation in IPS consistently had a significant unique contribution to explaining APS (supporting Hypothesis 1b), which demonstrates the relevance of crystallized strategic knowledge on generating and testing hypotheses (Greiff & Fischer, 2013a,b).

The unique contribution of solution in IPS was significant in the high-school student sample but not in the university-student sample (partially supporting Hypothesis 1c). In this regard we didn't expect to

Table 5

Unique and common commonality coefficients ΔR^2 and the corresponding percent of explained variance (%total) for each predictor in the regressions of GPA on representation, solution, APS, and reasoning based on the high-school student sample.

Variables	ΔR^2	%total
Unique to representation	0.019	10.73
Unique to solution	0.019	11.02
Unique to APS	0.040	23.34
Unique to reasoning	0.002	1.36
Common to representation, and solution	-0.013	-7.27
Common to representation, and APS	0.031	17.74
Common to solution, and APS	-0.013	-7.79
Common to representation, and reasoning	0.003	1.46
Common to solution, and reasoning	-0.000	-0.10
Common to APS, and reasoning	0.010	5.62
Common to representation, solution, and APS	0.020	11.80
Common to representation, solution, and reasoning	-0.001	-0.51
Common to representation, APS, and reasoning	0.017	10.12
Common to solution, APS, and reasoning	-0.002	-0.91
Common to representation, solution, APS, and reasoning	0.040	23.39
Total	0.172	100.00

Table 6

Unique and common commonality coefficients ΔR^2 and the corresponding percent of explained variance (%total) for each predictor in the regressions of GPA on representation, solution, APS, and reasoning based on the university student sample.

Variables	ΔR^2	%total
Unique to representation	0.006	3.33
Unique to solution	0.005	2.40
Unique to APS	0.135	72.59
Unique to reasoning	0.004	1.95
Common to representation, and solution	0.018	9.63
Common to representation, and APS	-0.002	-0.90
Common to solution, and APS	-0.003	-1.44
Common to representation, and reasoning	-0.001	-0.69
Common to solution, and reasoning	0.001	0.29
Common to APS, and reasoning	0.001	0.42
Common to representation, solution, and APS	0.013	6.93
Common to representation, solution, and reasoning	-0.001	-0.63
Common to representation, APS, and reasoning	0.001	0.53
Common to solution, APS, and reasoning	0.001	0.37
Common to representation, solution, APS, and reasoning	0.010	5.23
Total	0.186	100.00

find a unique contribution because of the close conceptual relation between solution and reasoning. However according to the theoretical framework of Fischer et al. (2012), in many cases of IPS it may be more effective to rely on implicit knowledge about inputs that work (instance-based knowledge) than to actually reason about structural knowledge (i.e., reasoning and representation). This kind of knowledge is not addressed in the causal model that indicates representation in IPS (Greiff & Fischer, 2013b). Future studies should elaborate on these incremental aspects of solution in IPS (cf. Fischer et al., 2012) in more detail in order to clarify what they depend upon.

4.2. Research question 2: can GPA be explained by unique aspects of APS or the facets of IPS over and above reasoning?

Our second research question was whether IPS and APS can explain school grades and if they have unique contributions compared to each other. Indeed, we found substantial correlations between Grade Point Average and both APS and the facets of IPS (supporting Hypothesis 2a). More importantly, APS had a unique contribution to explaining school grades beyond reasoning and both aspects of IPS in both samples (supporting Hypothesis 2d). This was expected because APS was known to be highly predictive for school achievements in different domains (OECD, 2004; Scherer & Tiemann, 2014), but it has never been proven empirically up to now.

With regard to representation and solution in IPS we found unique contributions in the high-school student sample but not in the university student sample (partially supporting Hypotheses 2b and 2c). With regard to representation in IPS we expected a unique contribution whereas for solution in IPS we did not expect a unique contribution because of its close conceptual relation to deductive reasoning (however, there are theoretical differences, see above). Again, future research is needed to determine the conditions for an incremental value of IPS with regard to the regression of GPA.

4.3. Shortcomings

First, comparisons between samples have to be drawn with caution, as the tests we used differed between samples.⁶ Of course they were highly similar in nature and can be assumed to address the same constructs. Nevertheless, we see a potential shortcoming here. Second, the operationalization of school grades may have had different meanings for high-school students (current grades) compared to university students (final Grade Point Average). Future studies should additionally

⁶ This was done because of the different levels of competence between the two samples.

investigate concurrent measures of success for university students, to replicate our findings and to further validate measures of IPS and APS (especially as skills and competence may change over time, see Molnár, Greiff, & Csapó, 2013). On the other hand, both shortcomings also highlight the robustness of our findings: Even in spite of these differences, most of our findings were consistent between samples highlighting the generalizability of our conclusions.

Please note, up to now the incremental value of assessing IPS strategies compared to reasoning has been demonstrated using a variety of reasoning tests and school grades (Greiff & Fischer, 2013a; Greiff et al., 2013b,c; Wüstenberg et al., 2012) and the results seem to be independent of different operationalizations as they can be attributed to a latent underlying construct (Greiff et al., 2013b). One may argue, that representation and solution in IPS do not cover the whole range of interactive strategies, but – even if we totally agree with this point – the high proportion of variance in APS explained by representation and solution beyond reasoning ($R^2 = .33$ to $.52$) indicates the centrality of the strategies assessed for cross-curricular Problem Solving Competency (which is basically the understanding of problem solving in PISA 2003 and PISA 2012).

4.4. Summary and outlook

Our findings clearly indicate close relations between Interactive and APS (beyond reasoning), as well as the incremental value of APS for predicting school grades (beyond IPS and reasoning) in both samples. For the first time, these findings show that APS – just like IPS – requires more than reasoning and thus is a promising measure of Problem Solving Competency for both High-School students and University Students. We did not find consistent evidence for a unique contribution of IPS (beyond APS and reasoning). More specifically, we found an incremental value of representation in IPS in the high-school sample, but not in the university student sample. This finding empirically supports the idea of complementing measures of APS with measures of IPS in school-related assessments, as it was done in PISA 2012 for instance.

In both samples, we found IPS to be closely related to APS beyond reasoning. Thus, both APS and IPS seem to address a common core of Problem Solving Competency (beyond reasoning) and APS seems to be more closely related to school grades than IPS. Our findings indicate that the analytic aspects of Problem Solving Competency assessed by IPS tasks (Scherer & Tiemann, 2014) may account for the incremental value of IPS beyond reasoning reported in previous studies on university student samples (e.g., Wüstenberg et al., 2012).

Implications for educational practice are manifold: In assessment contexts APS may be used to complement traditional assessment instruments (especially when GPA is an intended criterion) and in training contexts PSC maybe fostered by training and teaching strategic knowledge concerning how to acquire and how to apply knowledge or how to analyze evidence in IPS and APS (Scherer & Tiemann, 2014) – and much more easily so than reasoning. Realistic complex problems – from managing a Tailorshop (e.g., Danner et al., 2011) to solving in-basket tasks (e.g., Fischer & Funke, 2013) – shed light on the complexity of Problem Solving Competency itself (cf. Fischer et al., 2012). The tasks presented in the current study may be a first step to assessing certain aspects of PSC reliably, but the full potential of computer-based assessment still waits to be fully exploited.

Acknowledgments

This research was funded by grants from the German Research Foundation (DFG Az. Fu 17314-2) and from the German Federal Ministry of Education and Research (BMBF FKZ 01JG1062). We are grateful to the TBA group at DIPF (<http://tba.dipf.de>) for providing the MicroDYN authoring tool CBA Item Builder and technical support.

References

- Bassok, M., & Novick, L.R. (2012). Problem solving. In K.J. Holyoak, & R.G. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (pp. 413–432). New York, NY: Oxford University Press.
- Berry, D.C., & Broadbent, D.E. (1984). On the relationship between task performance and associated verbalizable knowledge. *The Quarterly Journal of Experimental Psychology*, 36, 209–231.
- Boggiano, A.K., Flink, C., Shields, A., Seelbach, A., & Barrett, M. (1993). Use of techniques promoting students' self-determination: Effects on students' analytic problem-solving skills. *Motivation and Emotion*, 17, 319–336.
- Carpenter, P.A., Just, M.A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, 97(3), 404–431.
- Cushing, J.M. (1977). Periodic time-dependent predator–prey systems. *SIAM Journal on Applied Mathematics*, 32, 82–95.
- Danner, D., Hagemann, D., Holt, D.V., Hager, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). Measuring performance in dynamic decision making: Reliability and validity of the tailorshop simulation. *Journal of Individual Differences*, 32, 225–233.
- Dörner, D. (1996). *The logic of failure: Recognizing and avoiding error in complex situations*. New York, NY: Perseus Books.
- Duncker, K. (1945). On problem-solving. *Psychological Monographs*, 58(5).
- Edwards, W. (1962). Dynamic decision theory and probabilistic information processings. *Human Factors*, 4(2), 59–74.
- Fischer, A., & Funke, J. (2013, August). Problem solving – More than the sum of its parts? Paper presented at the 15th Biennial Conference of the European Association for Research on Learning and Instruction (EARLI), Munich, Germany.
- Fischer, A., Greiff, S., & Funke, J. (2012). The process of solving complex problems. *Journal of Problem Solving*, 4, 19–42.
- Fischer, A., Greiff, S., & Funke, J. (2015). The history of complex problem solving. In B. Csapó, J. Funke, & A. Schleicher (Eds.), *The nature of problem solving*. Paris: OECD (under review).
- Fleischer, J., Buchwald, F., Wirth, J., Rumann, S., & Leutner, D. (2015). Analytical problem solving: Potentials and manifestations. In B. Csapó, J. Funke, & A. Schleicher (Eds.), *The nature of problem solving*. Paris: OECD (under review).
- Fleischer, J., Wirth, J., & Leutner, D. (2014). Effekte der kontextuellen Einkleidung von Testaufgaben auf die Schülerleistungen im analytischen Problemlösen und in der Mathematik [Effects of test context on students' achievements in analytical problem solving and mathematics]. *Zeitschrift für Pädagogische Psychologie*, 28, 217–227.
- Fleischer, J., Wirth, J., Rumann, S., & Leutner, D. (2010). Strukturen fächerübergreifender und fachlicher Problemlösekompetenz. Analyse von Aufgabenprofilen. Projekt Problemlösen [Structures of cross-curricular and subject-related problem solving competence – Analysis of task profiles]. *Zeitschrift für Pädagogik*, 56, 239–248 (Beihft).
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7, 69–89.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- Funke, J. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, 11, 133–142.
- Frischkorn, G. T., Greiff, S., & Wüstenberg, S. (2014). The development of complex problem solving in adolescence: A latent growth curve analysis. *Journal of Educational Psychology*, 106(4), 1007–1020.
- Gonzalez, C., Lerch, J.F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science*, 27, 591–635.
- Greiff, S. (2012). *Individualdiagnostik der komplexen Problemlösefähigkeit [Individual diagnostics of complex problem solving ability]*. Münster: Waxmann.
- Greiff, S., & Fischer, A. (2013a). Der Nutzen einer komplexen Problemlösekompetenz: Theoretische Überlegungen und empirische Befunde [Usefulness of complex problem solving competency: Theoretical considerations and empirical results]. *Zeitschrift für Pädagogische Psychologie*, 27, 27–39.
- Greiff, S., & Fischer, A. (2013b). Measuring complex problem solving: An educational application of psychological theories. *Journal for Educational Research Online*, 5(1), 38–58.
- Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015). Assessing complex problem-solving skills with multiple complex systems. *Thinking and Reasoning*. <http://dx.doi.org/10.1080/13546783.2014.989263> (in press).
- Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013b). A multitrait-multimethod study of assessment instruments for complex problem solving. *Intelligence*, 41, 579–596.
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013c). Complex problem solving in educational settings – Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, 105, 364–379.
- Heller, K.A., & Perleth, C. (2000). *KFT 4-12 + R: Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision [Cognitive ability test for grades 4 through 12, revised]*. Göttingen: Beltz Test.
- Horn, J., & Masunaga (2006). A merging theory of expertise and intelligence. In K.A. Ericsson, N. Charness, P.J. Feltovich, & R.R. Hoffman (Eds.), *Handbook of expertise and expert performance* (pp. 587–612). Cambridge: Cambridge University Press.
- Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.
- Klahr, D. (2000). Scientific discovery as problem solving. In D. Klahr (Ed.), *Exploring science. The cognition of development of discovery processes* (pp. 21–40). Cambridge: MIT Press.

- Leutner, D., Fleischer, J., & Wirth, J. (2006). Problemlösekompetenz als Prädiktor für zukünftige Kompetenz in Mathematik und in den Naturwissenschaften [Problem solving competence as a predictor of future competencies in mathematics and science]. In M. Prenzel (Eds.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (pp. 119–137). Münster: Waxmann.
- Leutner, D., Fleischer, J., Wirth, J., Greiff, S., & Funke, J. (2012). Analytische und dynamische Problemlösekompetenz im Lichte internationaler Schulleistungsvergleichsstudien: Untersuchungen zur Dimensionalität [Analytical and dynamic problem-solving competency from the perspective of studies on academic achievement in international comparison: Analysis of dimensionality]. *Psychologische Rundschau*, 63, 34–42. <http://dx.doi.org/10.1026/0033-3042/a000108>.
- Leutner, D., Funke, J., Klieme, E., & Wirth, J. (2005). Problemlösefähigkeit als fächerübergreifende Kompetenz [Problem-solving-competency as a cross-curricular key competency]. In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern* (pp. 7–20). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Leutner, D., Klieme, E., Meyer, K., & Wirth, J. (2004). Problemlösen. In PISA-Konsortium Deutschland (Ed.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (pp. 147–175). Münster: Waxmann.
- Leutner, D., Wirth, J., Klieme, E., & Funke, J. (2005). Ansätze zur Operationalisierung und deren Erprobung im Feldtest zu PISA 2000 [Approaches for operationalization and their evaluation in PISA 2000 field study]. In E. Klieme, D. Leutner, & J. Wirth (Eds.), *Problemlösekompetenz von Schülerinnen und Schülern* (pp. 21–36). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Liepmann, D. (2007). *Intelligenz-Struktur-Test 2000 R: IST 2000 R [Intelligence structure test 2000 (revised)]*. Göttingen: Hogrefe.
- Mayer, R.E. (2011). *Problem solving and reasoning*. Learning and cognition in education, 112–117.
- Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity*, 9, 35–45.
- Muthén, L.K., & Muthén, B.O. (2007). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Muthén, B.O., & Muthén, L.K. (2008). *Mplus (Version 5.1)*. Los Angeles, CA: Muthén & Muthén.
- Nimon, K., Lewis, M., Kane, R., & Haynes, R.M. (2008). An R package to compute commonality coefficients in the multiple regression case: An introduction to the package and a practical example. *Behavior Research Methods*, 40(2), 457–466.
- O'Brien, R.M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, 41(5), 673–690.
- Organisation for Economic Cooperation and Development (OECD) (2003). *The PISA 2003 assessment framework: Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.
- OECD (2014). *PISA 2012 results: Creative problem solving: Students' skills in tackling real-life problems (Volume V)*. Paris: OECD Publishing.
- Organisation for Economic Cooperation and Development (OECD) (2004b). *PISA 2012 Results: Creative problem solving: Students' skills in tackling real-life problems, Vol. V*. Paris: OECD.
- Postlethwaite, Bennett Eugene (2011). "Fluid ability, crystallized ability, and performance across multiple domains: a meta-analysis.". PhD thesis, University of Iowa <http://ir.uiowa.edu/etd/1255>.
- Raven, J. (2000). Psychometrics, cognitive ability, and occupational performance. *Review of Psychology*, 7(1–2), 51–74.
- Rya-Mukherjee, J., Nimon, K., Mukherjee, S., Morris, D.W., Slotow, R., & Hamer, M. (2014). Using commonality analysis in multiple regressions: A tool to decompose regression effects in the face of multicollinearity. *Methods in Ecology and Evolution*, 5(4), 320–328.
- Scherer, R., & Tiemann, R. (2014). Evidence on the effects of task interactivity and grade level on thinking skills involved in complex problem solving. *Thinking Skills and Creativity*, 11, 48–64.
- Schoppek, W., & Putz-Osterloh, W. (2003). Individuelle Unterschiede und die Bearbeitung komplexer Probleme [Individual differences and the solution of complex problems]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24(3), 163–173.
- Simon, H.A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, 7, 268–288.
- Strohschneider, S., & Guss, 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.
- Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, 26(2), 265–283. <http://dx.doi.org/10.1007/s10648-013-9243-1>.
- Verhulst, P.-F. (1839). Notice sur la loi que la population suit dans son accroissement. In J.G. Garnier, & A. Quételet (Eds.), *Correspondance Mathématique et Physique, Vol. 10* (pp. 113–121). Brussels (Belgium): Société Belge de Librairie.
- Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, 10, 329–345.
- Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving: More than Reasoning? *Intelligence*, 40, 1–14.