Intelligence 41 (2013) 289–305

Contents lists available at SciVerse ScienceDirect

Intelligence

Students' complex problem-solving abilities: Their structure and relations to reasoning ability and educational success

Philipp Sonnleitner [a,*], Ulrich Keller [a], Romain Martin [a], Martin Brunner [b]

[a] University of Luxembourg, Centre for Educational Measurement and Applied Cognitive Science (EMACS), Campus Kirchberg, 6, rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Luxembourg

[b] Free University of Berlin, ISQ, Berlin-Brandenburg Institute for School Quality Improvement, Otto-von-Simson-Str. 15, 14195 Berlin, Germany

ARTICLE INFO

Article history:
Received 21 October 2012
Received in revised form 17 April 2013
Accepted 9 May 2013
Available online xxxx

Keywords:
Complex problem solving
Reasoning
Educational success
Genetics Lab
Microworlds

ABSTRACT

Complex Problem Solving (CPS) is considered to be a promising candidate for capturing higher order thinking skills that are emphasized in new educational curricula but are not adequately measured by traditional intelligence tests. However, little is known about its psychometric structure and its exact relation to intelligence and educational success—especially in student populations. This study is among the first to use a large and representative sample of secondary school students ($N = 563$) to examine different measurement models of CPS—that conceptualize the construct as either faceted or hierarchical—and their implications for the construct's validity. Results showed that no matter which way it was conceptualized, CPS was substantially related to reasoning and to different indicators of educational success. Controlling for reasoning within a joint hierarchical measurement model, however, revealed that the impressive external validity was largely attributable to the variance that CPS shares with reasoning, suggesting that CPS has only negligible incremental validity over and above traditional intelligence scales. On the basis of these results, the value of assessing CPS within the educational context is discussed.

1. Introduction

Intelligence tests were originally developed for educational settings (Binet & Simon, 1905; Deary, Strand, Smith, & Fernandes, 2007) to predict whether a student would succeed in mastering academic subjects or not (Mayer, 2000). Therefore, virtually all intelligence tests include subtests to capture students' abilities to solve problems and to reason. Although intelligence tests have not changed much since their invention more than 100 years ago (Hunt, 2011; Sternberg & Kaufman, 1996; Sternberg, Lautrey, & Lubart, 2003), they still fulfill this purpose quite well (Deary, 2012; Hunt, 2011; Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012; Naglieri & Bornstein, 2003).

In recent years, however, educational systems have been in a transition caused mainly by dramatic innovations in information technology (IT). This has produced a significant change in the student population; thus, today's students are described as "digital natives" (Prensky, 2001) or the "net generation" (Tapscott, 1998). Moreover, new educational goals that focus on students' problem-solving abilities have been set (Bennett, Jenkins, Persky, & Weiss, 2003; Ridgway & McCusker, 2003). The integration of such higher order thinking skills (Kuhn, 2009) in educational curricula is necessary to prepare students to solve the complex problems of today's world. Consequently, large-scale assessments such as the Program for International Student Assessment (PISA) have extended their evaluation scheme of key outcomes of the educational system to include problem-solving abilities (see Leutner, Fleischer, Wirth, Greiff, & Funke, 2012; Wirth & Klieme, 2003).

Given these changes, we must ask whether traditional intelligence tests (which include subtests that assess students' abilities to reason and solve problems) still capture the cognitive skills that are vital for success in education or whether the

* Corresponding author. Tel.: +352 466644 9514.
E-mail addresses: philipp.sonnleitner@uni.lu (P. Sonnleitner), ulrich.keller@uni.lu (U. Keller), romain.martin@uni.lu (R. Martin), martin.brunner@isq-bb.de (M. Brunner).

0160-2896/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.intell.2013.05.002

assessment of new constructs would be more beneficial. According to many authors, a potential alternative in this context is the construct *complex problem solving* (*CPS*; Funke, 2010). CPS is measured with computer-based problem-solving scenarios (also known as microworlds), which receive high acceptance among today's students (Ridgway & McCusker, 2003; Sonnleitner et al., 2012). These microworlds provide detailed information about students' problem-solving behaviors and strategies when they address complex and dynamic problems (Fischer, Greiff, & Funke, 2011; Funke, 2001).

The results obtained with CPS measures are promising: not only that CPS has been found to be substantially correlated with intelligence (thus indicating that CPS captures central cognitive abilities that are similar to those captured by intelligence tests); recently, studies have also reported that CPS explains individual differences in external criteria over and above what is accounted for by intelligence. Measures of CPS were found to possess incremental validity beyond intelligence tests in predicting supervisor ratings (Danner, Hagemann, Schankin, Hager, & Funke, 2011) and—crucial to the educational context— grade point average (Wüstenberg, Greiff, & Funke, 2012).

In sum, there is promising evidence that CPS may be a reliable and valid representation of higher order thinking skills and problem-solving behavior. However, most previous research on CPS has been based on (highly) selected samples of university students (e.g., psychology students). Empirical studies based on samples of secondary school students are still rare. Nevertheless, insights obtained for this population are vital for evaluating whether tests that measure CPS can potentially be used to assess students' higher order thinking skills because these skills are required by today's educational curricula. The major goal of the present paper was to significantly contribute to the knowledge about CPS by using a large heterogeneous sample of secondary school students. In doing so, we analyzed (a) the structure of CPS (as reflected by corresponding measurement models), (b) its relation to reasoning, (c) its ability to predict success in school (external validity), and (d) its ability to predict students' educational success over and above traditional intelligence tests (incremental validity). To this end, we put special emphasis on the (joint) hierarchical structure of intelligence and CPS. This allowed us to disentangle the individual differences in students' higher order thinking skills that are shared between traditional intelligence tests and measures of CPS from the individual differences that are unique to each of these measures. In doing so, the present results provide important insights into whether CPS is able to provide information that can be used to predict students' educational success over and above intelligence or not.

2. The relations between CPS and intelligence on latent and manifest levels

Cognitive abilities are not directly observable; rather, they are latent constructs. One core idea that is used in the assessment process is that these latent constructs are considered to be distinct from their manifest measures. This distinction emphasizes the critical importance of the measurement model, which links latent variables to their corresponding measures (Bollen & Lennox, 1991; Borsboom, Mellenbergh, & Van Heerden, 2003; Brunner, Nagy, & Wilhelm, 2012; Edwards & Bagozzi, 2000). Crucially, it has been shown that the choice of the measurement model may have severe consequences and even lead to different results when the relation between constructs is under investigation (Brunner, 2008; Hornung, Brunner, Reuter, & Martin, 2011). Thus, when the relation between CPS and intelligence is under investigation, theoretical considerations as well as former empirical results are vital for ensuring that well-grounded measurement models of both constructs are designed to represent their structures.

2.1. The structure of complex problem solving

Up to now, there has been no widely accepted definition of the latent construct of CPS (Fischer et al., 2011; Frensch & Funke, 1995; Quesada, Kintsch, & Gomez, 2005). However, most definitions encompass the ability to overcome barriers in order to achieve a target state within a complex and dynamically changing environment (e.g., Buchner, 1995; Fischer et al., 2011; Frensch & Funke, 1995; Mayer & Wittrock, 1996). However, greater consensus exists with regard to the measurement of CPS by means of computer-based scenarios, so-called microworlds that should mirror complex problems. Microworlds that are used in CPS research were originally intended to overcome the limitations of traditional intelligence tests (Funke, 1993) and are described by several key characteristics: They (a) consist of several variables that (b) are highly interconnected and (c) change over time (i.e., are dynamic). Crucially, (d) these underlying connections are not transparent, and (e) the test taker has to achieve several partly contradictory goals (Funke, 2001, 2003, 2010). Notably, some of these characteristics are not shared with intelligence tests (e.g., on intelligence tests, tasks do not change over time or do not require the achievement of multiple goals). Typically, test takers interact with microworlds in two phases. In the first phase, they manipulate the microworld's variables in order to acquire knowledge. This knowledge must then be applied in a second phase in order to achieve several goals. An example of a contemporary microworld is the Genetics Lab (GL; Sonnleitner et al., 2012) shown in Fig. 1. In the GL, three scores are obtained, reflecting the test taker's ability to (a) retrieve information about the problem by applying an appropriate exploration strategy, (b) build a correct mental model of the problem, and (c) apply the gathered knowledge to achieve certain problem states. These abilities are described as central facets of CPS (and also show large overlap with the abilities that are tested by typical tests of reasoning ability; see also Fischer et al., 2011; Wüstenberg et al., 2012).

Studies investigating the structure of CPS have provided mixed results. Kröner, Plass, and Leutner (2005) were among the first to report three different facets of CPS that could be empirically distinguished. These facets corresponded to the typically obtained scores in such scenarios (see above) and were described as *rule identification*, referring to the quality of the applied exploration strategy, *rule knowledge*, and *rule application*. The study reported by Kröner et al. (2005) exclusively included students from high school (German