Learning and Individual Differences 29 (2014) 18–29

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Cross-national gender differences in complex problem solving and their determinants

Sascha Wüstenberg [a,*], Samuel Greiff [a], Gyöngyvér Molnár [b], Joachim Funke [c]

[a] University of Luxembourg, Luxembourg
[b] University of Szeged, Hungary
[c] University of Heidelberg, Germany

ARTICLE INFO

Article history:
Received 20 July 2012
Received in revised form 18 July 2013
Accepted 11 October 2013

Keywords:
Complex problem solving
MicroDYN
Measurement invariance
Gender differences
Nationality differences

ABSTRACT

The present study examined cross-national gender differences in domain-general complex problem solving (CPS) and their determinants. A CPS test relying on the MicroDYN approach was applied to a sample of 890 Hungarian and German high school students attending 8th to 11th grade. Results based on multi-group confirmatory factor analyses showed that measurement invariance of CPS was found across gender and nationality. Analyses of latent mean differences revealed that males outperformed females and German students outperformed Hungarian students. However, these results were caused by Hungarian females performing worse than all other groups. Further analyses of logfiles capturing process data of the interaction of participants with the task showed that Hungarian females less often used vary-one-thing-at-a-time strategy, which lead to considerably worse knowledge acquisition. Results imply that analyzing process data such as use of strategies is highly advisable to identify determinants of overall performance differences in CPS across groups of interest.

1. Introduction

Over the last decades, reports on individual differences in students' performance across gender or nationality have strongly influenced educational policies. For instance, results of the Programme for International Student Assessment (PISA) 2000 led to changes of the educational system and revisions of educational standards in Germany (Wernstedt & John-Ohnesorg, 2009), because German students underperformed in comparison to neighboring countries. Especially performance differences in domain-specific areas such as mathematical ability play an important role not only in educational research (Else-Quest, Hyde, & Linn, 2010; Lindberg, Hyde, Petersen, & Linn, 2010), but also in high stakes assessments such as Trends in International Mathematical and Science Study (TIMSS) or PISA.

However, only little is known about individual differences in students' domain-general competencies notwithstanding an increasing scientific and public interest. For instance, domain-general problem solving competency was assessed in the 2012 cycle of PISA, which was conducted by the Organisation for Economic Co-operation and Development (OECD) with results scheduled for publication in 2014. More specifically, the OECD emphasizes the high educational and socio-economical relevance of domain-general problem solving in everyday life as it "provides a basis for future learning" (OECD, 2010, p. 7). Thus, domain-general problem solving is considered a highly relevant competency for students that should be developed in addition to domain-specific knowledge in school subjects. Within domain-general problem solving, (non-interactive) analytical problem solving and (interactive) complex problem solving (CPS) can be distinguished as subordinate constructs (Fischer, Greiff, & Funke, 2012; OECD, 2010). Whereas analytical problem solving is usually measured with static paper–pencil tasks, complex problem solving (CPS)[1] includes tasks enabling interactions between user and task situation (Wirth & Klieme, 2003). Recently, the OECD emphasized the importance of the domain-general competency to interactively deal with novel problems:

> *Mobilisation of prior knowledge is not sufficient to solve novel problems in many everyday situations. Gaps in knowledge must be filled by observation and exploration of the problem situation. This often involves interaction with a new system to discover rules that in turn must be applied to solve the problem.*
>
> [OECD, 2010, p. 15]

* Corresponding author at: University of Luxembourg, 6, rue Richard Coudenhove Kalergi, 1359 Luxembourg-Kirchberg, Luxembourg. Tel.: +352 466644 9704.
E-mail address: Sascha.Wuestenberg@uni.lu (S. Wüstenberg).

[1] The OECD used the term interactive problem solving (OECD, 2010) instead of CPS, referring to the interactive nature of the task. In the present paper, we use the term complex problem solving (CPS), which emphasizes the aspect of the underlying system's complexity. Both terms are used synonymously, but CPS is most established in research (Dörner, 1986, 1990; Funke, 2001, 2010).

1041-6080/$ – see front matter
http://dx.doi.org/10.1016/j.lindif.2013.10.006

CPS tasks as described in this quotation usually contain many highly interrelated elements and system states of the tasks change dynamically (cf. Fischer et al., 2012; Funke, 2001). By interacting with CPS tasks, problem solvers have to overcome barriers between a given initial state and a goal state (Funke, 2012; Mayer, 2003). Thereby, they explore and integrate information to discover rules that must be applied to solve the problem (Buchner, 1995). CPS tasks are applied fully computer-based (Wirth & Klieme, 2003), giving researchers the opportunity to not only evaluate outcomes (e.g., whether a problem is solved or not), but also to analyze process data (e.g., how a problem solver interacts with a problem). This enables analyses of determinants of performance, for instance, which strategies are used to gather information and to solve a certain problem.

While interacting with the task, problem solvers (1) build a problem representation and (2) derive a problem solution (Novick & Bassok, 2005). These two major components of problem solving are usually measured by two dimensions: the competency of problem solvers to gain new knowledge during the interaction with the task — (1) knowledge acquisition — and to apply that knowledge to solve the task, (2) knowledge application (Bühner, Kröner, & Ziegler, 2008; Funke, 2001).

Recently conducted studies show that both dimensions knowledge acquisition and knowledge application can be empirically distinguished in domain-general CPS research (Bühner et al., 2008; Greiff, Wüstenberg, & Funke, 2012; Wüstenberg, Greiff, & Funke, 2012). Furthermore, CPS predicts supervisor ratings of participants' overall job performance (Danner, Hagemann, Schankin, Hager, & Funke, 2011) and school grade point average (Greiff & Fischer, 2013; Wüstenberg et al., 2012) even beyond reasoning. However, to our knowledge, no studies have yet been conducted analyzing individual differences in students' CPS performance and their determinants with regard to gender and nationality.

As a prerequisite for analyses and interpretations of overall differences in CPS performance and their determinants, it has to be ensured that there are no systematic demographic subgroup biases. For instance, in educational research on students' abilities as well as in intelligence testing, plenty of research has been conducted to ensure that measurement devices allow an unbiased measurement of the construct of interest across subgroups (e.g., Bowden, Saklofske, & Weiss, 2011; Chen, 2012; Gardner & Qualter, 2011). Especially performance differences with regard to gender and nationality often raise interest and concerns (cf. Else-Quest et al., 2010), leading to extensive discussions about determinants of performance, and – as outlined above in case of differing performance across nationality – to changes of whole educational systems (Wernstedt & John-Ohnesorg, 2009). Thus, it is vitally important to understand whether between-group differences in cognitive performance with regard to gender and nationality reflect true differences in the construct of interest, or different psychometric properties of the underlying measurement scale (Brown, 2006). But even if performance differences are valid, educationalists are not only interested in knowing *that* performance differences exists, but they also want to know *why* they exist in order to be able to foster the underlying competency by applying appropriate interventions. With regard to CPS, research on accurate measurement of performance differences across groups as well as determinants of performance is still in its infancy.

As will be further outlined, it is yet unclear whether CPS can be measured with equal validity across (1) gender or (2) nationality and analyses on individual differences in CPS performance are scarce. In fact, joint analyses of differences including (3) both gender and nationalities are non-existent. Particularly the latter is of high interest, because if gender differences vary in specific countries more than in others, cross-national patterns may reflect "inequities in educational and economic opportunities" regarding gender (Else-Quest et al., 2010, p.103). There are also only few studies, which (4) investigate determinants of performance differences in CPS by analyzing process data gathered while participants interact with the task environment.

To this end, based on a sample of Hungarian and German high school students, (1) we will evaluate whether CPS can be measured with equal validity across gender and investigate gender differences in mean CPS performance. (2) We will analogously evaluate whether CPS can be measured with equal validity across Germans and Hungarians and investigate differences in mean CPS performance. (3) Further, we conduct combined analyses to study interaction effects of gender and nationality. Therefore, our sample is separated in four groups containing German males, German females, Hungarian males, and Hungarian females to evaluate whether CPS can be measured with equal validity across gender and nationality and to investigate mean differences across four groups. (4) Finally, we investigate determinants of mean performance differences across groups in knowledge acquisition by analyzing process data including behavioral patterns of participants gathered during their exploration of the tasks.

1.1. Measurement invariance and latent mean differences across gender

As a prerequisite of interpreting gender differences in CPS, structural stability of the construct has to be secured by evaluating measurement invariance (cf. Byrne & Stewart, 2006; Sass, 2011), a state of the art procedure frequently applied for measures of cognitive performance (e.g., mathematical ability; Brunner, Krauss, & Kunter, 2008). For instance, it was shown that the factor structure of the Wechsler Intelligence Scale for Children (WISC) does not change across gender (Chen & Zhu, 2008). However, although various CPS measures exist (e.g., *Genetics Lab*, Sonnleitner et al., 2012; *MultiFlux*, Kröner, Plass, & Leutner, 2005; *NewFire*, Rigas, Carling, & Brehmer, 2002; and *Tailorshop*, Süß, 1996), no studies have been conducted analyzing measurement invariance with regard to gender. Only recently, it was shown that CPS can be measured invariant across Hungarian high school students in different grades (Greiff et al., 2013).

With regard to gender differences in CPS, previous findings are contrary to results on reasoning ability, in which reported gender differences slightly favor females showing rather small or marginal effect sizes (Brunner et al., 2008; Halpern & LaMay, 2000; Jensen, 1998). In CPS, only few studies investigated gender differences, pointing towards a considerable advantage of males (Jensen & Brehmer, 2003; Wittmann & Hattrup, 2004; Wittmann & Süß, 1999). However, the study of Jensen and Brehmer (2003) was based on a very small sample with limited generalizablity ($N = 15$; four males). As Wittmann and Hattrup (2004) integrated findings of Wittmann and Süß (1999) and two additional studies, we therefore only describe results of Wittmann and Hattrup (2004) in more detail.

Specifically, Wittmann and Hattrup (2004) pooled data of three independent studies using the CPS scenario *Tailorshop* (cf. Süß, 1996), in which participants have to maximize the company value of a tailor manufactory by controlling variables such as *number of workers* or *marketing*. In *Tailorshop*, investments in marketing have strong effects on the variable "demand", which in turn increases sales, being highly relevant for good performance within the simulation (Wittmann & Hattrup, 2004, p. 405). The authors showed that males outperformed females (Cohens' $d = .70$) and explained these differences by a higher level of risk aversiveness in females, who invested significantly less in marketing (i.e., varied the variable marketing to a lesser degree) compared to males. However, there are two other possible explanations than a lower amount of risk aversiveness in females not discussed by Wittmann and Hattrup (2004): (1) Males may rely on more efficient strategies while dealing with CPS tasks or (2) scenario effects may lead to males' better performance.

(1) In cognitive psychology, the use of strategies is known as (implicit) procedural knowledge (knowing how), which has to be applied in