Intelligence 41 (2013) 378–395

Contents lists available at ScienceDirect

Intelligence

Gender differences in the mean level, variability, and profile shape of student achievement: Results from 41 countries

Martin Brunner [a,*], Katarzyna Marta Gogol [b], Philipp Sonnleitner [b], Ulrich Keller [b], Stefan Krauss [c], Franzis Preckel [d]

[a] Free University of Berlin, Berlin–Brandenburg Institute for School Quality, Berlin, Germany
[b] University of Luxembourg, Luxembourg, Luxembourg
[c] University of Regensburg, Regensburg, Germany
[d] University of Trier, Trier, Germany

ARTICLE INFO

Article history:
Received 7 December 2012
Received in revised form 27 May 2013
Accepted 28 May 2013
Available online 29 June 2013

Keywords:
Student achievement
Cognitive gender differences
Large-scale assessment
Hierarchical factor models

ABSTRACT

A domain-specific hierarchical conceptualization of mathematics achievement can be represented by the standard psychometric model in which a single latent dimension accounts for observed individual differences in scores on the respective subdomains (e.g., quantity). Alternatively, a fully hierarchical conceptualization of achievement can be represented by a nested-factor model in which individual differences in subdomain-specific scores are explained by both general student achievement and specific mathematics achievement. The authors applied both models to study the gender similarity hypothesis, the greater male variability hypothesis, and the masking hypothesis, which predicts that gender differences in general student achievement mask gender differences in both the means and the variability of specific mathematics achievement. Representative data were obtained from 275,369 15-year-old students in 41 countries. The results supported these hypotheses in most countries, demonstrating that a fully hierarchical conceptualization of achievement in terms of the nested-factor model significantly contributes to a better understanding of gender differences in the mean level, variability, and shape of students' achievement profiles.

1. Introduction

Professionals working in the STEM fields (i.e., science, technology, engineering, and mathematics) drive the scientific and technological development of a country and thus contribute substantively to its economic growth and competitiveness, homeland security, and well-being (Halpern et al., 2007). Despite an increasing demand from the labor market for qualified employees in these fields (Committee on Equal Opportunities in Science and Engineering, 2000, Section 5), women are currently underrepresented in several STEM fields in most countries—particularly engineering and science (Halpern et al., 2007). In 2009, for example, women represented 56% of university students across the 27 member states of the European Union, but only 38% of the students in science, mathematics, and computing and 25% of those in engineering, manufacturing, and construction (Eurostat, 2011). Thus, to increase gender equity and to better exploit the female talent pool in meeting the increasing demand for highly qualified STEM professionals, research needs to identify the individual characteristics on which males and females differ.

It seems plausible that gender differences in students' mathematics achievement are a major factor in explaining the gender gap in STEM fields: Mathematics proficiency is an essential component of professional expertise in these fields. However, students differ not only in their mathematics achievement, but also in many other achievement domains such as reading and science. It is not only the level of

* Corresponding author at: Berlin–Brandenburg Institute for School Quality, Otto-von-Simson-Strasse 15, 14195 Berlin, Germany.
E-mail address: 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.2013.05.009

mathematics achievement (along with motivational aspects such as vocational interests) that affects choice, commitment, and success in educational and occupational careers. The pattern or profile of students' achievement across domains also substantively influences these outcomes (Lubinski & Benbow, 2006; Park, Lubinski, & Benbow, 2007). For example, highly gifted students with a specific strength in mathematics at age 13 (i.e., whose SAT-M scores are considerably larger than their SAT-V scores) have been found to be more likely to select mathematics courses as their favorite courses in high school and college, secure educational degrees and tenure-track positions in STEM fields, and secure a patent up to 20 years later (Lubinski & Benbow, 2006; Park et al., 2007). Moreover, drawing on large propensity samples of adolescent students, Brunner and colleagues (Brunner, 2008; Brunner, Lüdtke, & Trautwein, 2008) have found that specific strengths in mathematics are significantly related to a stronger interest in mathematics and higher levels of the mathematics self-concept but to weaker interests and lower levels of self-concepts in verbal subjects. In line with the findings for highly gifted students, these findings obtained for the general student population suggest that specific strengths in mathematics are associated with the gravitation of students to the STEM fields because of their achievement profiles and closely related configurations of achievement motivation.

Taken together, empirical studies have shown that specific strengths in mathematics in a student's achievement profile affect choice, commitment, and success in educational and occupational careers. However, despite the importance of the pattern of achievement profiles in general and specific strengths in mathematics in particular, most empirical research on gender differences in mathematics has focused exclusively on mathematics achievement and related subdomains (e.g., algebra, arithmetic, or geometry). The key objective of this article is to significantly contribute to this line of research by investigating how students' mathematics achievements are embedded in their broader achievement profiles. More specifically, drawing on representative data from 41 countries, we examined gender differences in the mean level, variability, and shape of students' achievement profiles. We focus our theoretical review and empirical analyses on adolescent students—that is, students at an age at which important decisions are made about future educational and occupational careers.

2. Conceptual and psychometric issues in the study of gender differences in achievement

2.1. Hierarchical perspectives on student achievement

2.1.1. The domain-specific hierarchical conceptualization and the standard model

For the purposes of the present study, we define *student achievement* as the knowledge and skills acquired through long-term, cumulative, and domain-specific learning processes (Baumert, Lüdtke, Trautwein, & Brunner, 2009). To account for the specificity of the knowledge and skills acquired within domains, most current theories on the structure of student achievement apply a *domain-specific hierarchical conceptualization*. For example, the assessment framework used in the Programme for International Student Assessment (PISA; Organisation for Economic Co-operation and Development [OECD], 2004) distinguishes four subdomains of mathematics (Fig. 1a): space and shape, uncertainty, quantity, and change and relationships. Two aspects of the domain-specific hierarchical conceptualization merit consideration. First, the other achievement domains assessed in PISA (i.e., reading, science, and problem solving) are also conceived as being hierarchically organized (Fig. 1a). Second, when the focus of research is on mathematics achievement, the mathematics section of the domain-specific hierarchical conceptualization of student achievement (see Fig. 1a) is typically studied in isolation.

When the research goal is to empirically investigate theories on the structure of achievement, such theories (or assessment frameworks) need to be translated into appropriate psychometric models. To this end, most researchers draw on a standard psychometric model. Fig. 1b shows the standard model for mathematics achievement. We call this model the *standard model* because it is the predominant model in educational and psychological research (Jarvis, MacKenzie, & Podsakoff, 2003). In the standard model applied in this article, we focus on the upper part of the hierarchy within the mathematics domain. Scores on individual test items are therefore summarized into subdomain-specific performance scores (see Bagozzi & Edwards, 1998, for a detailed description of this methodological approach). These performance scores are represented as manifest variables in the model (depicted as rectangles in Fig. 1b). *Mathematics achievement* is conceptualized as a *single* latent variable (depicted as an ellipse in Fig. 1b) that affects corresponding subdomain-specific performance scores. Individual differences in mathematics achievement are thus predicted to be (largely) responsible for individual differences observed in these subdomain-specific performance scores. The latent variables pointing to the subdomain-specific scores (i.e., $Sp_{specific}$, $C_{specific}$, $U_{specific}$, and $Q_{specific}$) represent the combined influence of reliable subdomain-specific knowledge and skills and random measurement error. Crucially, the domain-specific hierarchical conceptualization does not explicitly address the relationships between achievement domains; this is indicated by the dashed line that connects the various domains in Fig. 1a. To study these relationships, the standard model may be extended to cover several interrelated achievement domains such that each domain is represented by a single latent variable that is captured by several subdomain-specific performance scores. However, within the domain-specific conceptualization and the corresponding extended standard model, any empirical relationships between domains are typically taken for granted and are *not* explained further.

2.1.2. The fully hierarchical conceptualization and the nested-factor model

By contrast, the *fully hierarchical conceptualization of student achievement* (Fig. 1c) takes into account the substantial interrelations typically observed between domain-specific constructs. This conceptualization is based on current thinking about the structure of intelligence (McGrew, 2009), according to which, cognitive abilities (typically measured by intelligence tests) are fully hierarchically structured: General cognitive ability operates at the apex of the hierarchy; abilities specific to content domains (e.g., verbal ability) or cognitive processes