Studies in Educational Evaluation 71 (2021) 101069

Contents lists available at ScienceDirect

Studies in Educational Evaluation

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

Separating PIAAC competencies from general cognitive skills: A dimensionality and explanatory analysis

Lena Engelhardt[a,*], Frank Goldhammer[a,b], Oliver Lüdtke[b,c], Olaf Köller[c], Jürgen Baumert[d], Claus H. Carstensen[e]

[a] DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany
[b] Centre for International Student Assessment (ZIB), Germany
[c] Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany
[d] Max Planck Institute for Human Development, Berlin, Germany
[e] Otto-Friedrich University Bamberg, Bamberg, Germany

ARTICLE INFO

Keywords:
PIAAC literacy
PIAAC numeracy
General cognitive skills
Validity
Large-scale assessments

ABSTRACT

This study aims to investigate how test scores from PIAAC (Programme for the International Assessment of Adult Competencies) can be interpreted, by comparing the PIAAC competencies literacy and numeracy to reasoning and perceptual speed. Dimensionality analyses supported, that the PIAAC competencies can be separated into a common factor overlapping with reasoning and perceptual speed, and domain-specific factors. For the common and specific factors, relations to other variables were analyzed. The nested factor for PIAAC literacy was as expected unrelated to age, positively related to learning opportunities during one's lifetime, and positively related to literacy skill use. The nested factor for PIAAC numeracy was also as expected unrelated to age, against expectation unrelated to learning opportunities during one's lifetime, and as expected positively related to numeracy skill use. Results support the validity of the intended test score interpretation for PIAAC literacy, while results for PIAAC numeracy were less clear.

1. Introduction

International studies, such as the Programme for the International Assessment of Adult Competencies (PIAAC; OECD, 2016a), focus on the assessment of competencies, such as literacy and numeracy, among others. Competencies are important for competent handling of everyday situations. They are assumed to represent outcomes of educational processes that may occur inside and outside of schools (OECD, 2016a; Prenzel, Walter, & Frey, 2007). Hence, such competencies should be learnable, and educational settings offer learning opportunities for their acquisition. In turn, educational processes and systems can be evaluated by the assessment of learners' competencies (Klieme & Leutner, 2006; Klieme, Hartig, & Rauch, 2008). International studies were designed exactly for this very purpose, namely, to inform policy makers about strengths and weaknesses of the educational system (OECD, 2016a, p.16).

A central critique on large-scale assessments such as PIAAC or the OECD Programme for International Student Assessment (PISA; OECD, 2016b) concerns the cognitive test constructs (Hopfenbeck et al., 2018). What the tests actually assess was addressed in several studies on student data (Brunner, 2005; Nagy, 2006; Saß, Kampa, & Köller, 2017; see for a review: Baumert, Lüdtke, Trautwein, & Brunner, 2009). These results can, however, not be simply transferred to the adult population, targeted in the PIAAC study. The goal of the present study is therefore to investigate how test scores from PIAAC literacy and PIAAC numeracy can be interpreted by focusing on differences to test scores from reasoning and processing speed. The meaning of test scores is established by considering two sources of validity evidence (American Educational Research Association, American Psychological Association & National Council on Measurement in Education [AERA, APA, & NCME], 2014): *Evidence based on the internal structure* and *evidence based on relations to other variables*.

2. PIAAC competencies, reasoning, and processing speed

2.1. PIAAC competencies

PIAAC competencies are understood as "key information-processing

* Corresponding author.
E-mail addresses: engelhardtl@dipf.de (L. Engelhardt), goldhammer@dipf.de (F. Goldhammer).

https://doi.org/10.1016/j.stueduc.2021.101069
Received 24 February 2021; Received in revised form 28 June 2021; Accepted 12 August 2021
Available online 21 September 2021

skills" (OECD, 2016a, p. 16) that should enable respondents to participate in society and to master everyday life situations. Because success in everyday situations can depend on a variety of involved skills and cognitive processes (Klieme et al., 2008), competencies are characterized by complexity. Therefore, assessments confront test takers with contextualized tasks (cf. OECD, 2016a). In contrast, assessments of general cognitive skills such as reasoning (e.g. matrices tasks; Raven, 2000), ask for more general and abstract skills. PIAAC literacy tasks require for instance reading. Reading in turn is based on various processes such as word identification, construction of word meaning, short- and longtime memory, but also on working memory and drawing inferences (Kintsch, 1998). Hence, competence tasks require besides skills that are very specific to the task (e.g. word identification), also general skills (e.g. working memory). This is why overlaps between these two types of assessments – competencies and general cognitive skills – apparently exist and the added value of assessing competencies as it is done in large-scale assessment studies is discussed (Brunner, 2005; Nagy, 2006; Rindermann & Baumeister, 2015; Rindermann, 2006, 2007; Saß et al., 2017).

PIAAC competencies are assumed *to be learnable and to be based also on knowledge* (OECD, 2016a, p.16). Whether this intended interpretation of PIAAC competencies holds is investigated in the present study. Given the complexity of competencies and the various skills that are involved in the accomplishment of competence tests, as well as overlaps with general cognitive skills, the validity of the test score interpretation should be investigated by focusing on differences between PIAAC competencies and two general cognitive skills: reasoning and processing speed.

2.2. Reasoning and processing speed

Reasoning and processing speed are considered as components of general cognitive skills and are also targeted in classical ability theories (Carroll, 1993; Cattell, 1963; Horn, 1965; see for an overview McGrew, 2009). Classical ability theories distinguish fluid and crystallized skills. Research on brain structures and functions supports this classical distinction into fluid and crystallized skills (Nisbett et al., 2012; Waltz et al., 1999). Fluid skills, such as reasoning, are required to solve novel problems by drawing inferences or identifying relations. They are not based on acquired knowledge but on elementary processes such as processing speed (Schweizer & Koch, 2002). Processing speed, such as perceptual speed, describes the skill to perform easy tasks which contain over-learned elements automatically (McGrew, 2009). While fluid skills require controlled mental operations, these elementary processes can be performed automatically. Crystallized skills describe acquired knowledge, for instance in terms of language, and the application of this knowledge. When it comes to assessment, crystallized skills are assumed to be best assessed when specific knowledge needs to be applied, and fluid skills are best assessed when knowledge provides no advantage for performance tests (Cattell, 1963, p. 3), for instance when tests contain either novel or overlearned elements (Horn & Cattell, 1966, p. 255). In contrast to reasoning and processing speed, the PIAAC competencies literacy and numeracy are understood in the present study as containing also crystallized aspects as they are assumed to be learnable and based on acquired knowledge. This classical differentiation into fluid and crystallized skills provides a conceptual frame for the predictions in the present study.

3. Validation approach for PIAAC competence test scores

The present study collects validity evidence for the intended test score interpretation by investigating sources of validity evidence for PIAAC competencies always in comparison to the constructs reasoning and processing speed. This approach is chosen because PIAAC competence tasks require various skills, skills specific to the domain such as reading/numerical skills, but also general, domain-independent, skills. It is assumed that the constructs reasoning and processing speed differ from PIAAC competencies in aspects that are central to the intended test score interpretation described above. They serve therefore as comparison to derive conclusions regarding the validity of the intended test score interpretation. The goal is to investigate whether scores from PIAAC competencies tests can be interpreted as representing *domain-specific skills that can be acquired*. *Dimensionality analyses* are used to separate domain-specific aspects in PIAAC competencies from general aspects and contribute to *validity evidence based on the internal structure* (AERA, APA, & NCME, 2014). The interpretation of the separated aspects – as being domain-specific – should be validated by relations to other variables for which clear expectations exist based on theoretical assumptions and empirical findings. If the from reasoning and processing speed separated aspects fulfill the expected relation, *validity evidence based on relations to other variables* is collected for the intended test score interpretation (AERA, APA, & NCME, 2014) of the PIAAC competencies. The focus in the present study is on construct interpretation, the question on whether differences in test scores of the PIAAC competence tests literacy and numeracy can be interpreted as differences in the targeted construct.

3.1. Dimensionality

Already classical ability theories used hierarchical models for different mental abilities, with a common factor which captures what all have in common, and more specific factors at the lower level (see Carroll, 1993; Vernon, 1965). Since competence constructs and general cognitive skills are assumed to overlap, such an approach is chosen in several studies to investigate the validity of test score interpretations of competence constructs assessed in large-scale assessments. In those studies, common aspects between competence items beyond the items they were compared to – often reasoning items – are interpreted as domain-specific factors. For student data, such domain-specific factors were found (Baumert et al., 2009; Brunner, 2005; Nagy, 2006; Saß et al., 2017). However, previous results on student data cannot necessarily be transferred to older populations or to the competence tests used in PIAAC.

3.2. Relations to other variables

3.2.1. Age decline

Various studies investigate the relation of cognitive abilities to aging. Across studies, speed was found to be related to a strong age decline (Hartshorne & Germine, 2015; Lang, Weiss, Stocker, & von Rosenbladt, 2007; McArdle, Hamagami, Meredith, & Bradway, 2000; Verhaeghen & Salthouse, 1997). Age decline is also reported for reasoning (Horn & Cattell, 1967; Verhaeghen & Salthouse, 1997), while knowledge-based skills such as vocabulary knowledge declined very late or were even unrelated to age (Hartshorne & Germine, 2015; Horn & Cattell, 1967; Lang et al., 2007). More general predictions regarding the decline of abilities are made by classical ability theories, in which fluid aspects are assumed to decline with age to stronger extent than crystallized aspects (cf. Horn & Cattell, 1967). More recent literature supports that fluid reasoning, crystallized abilities, and also processing speed have different growth patterns (Li et al., 2004; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002). Supporting this distinction, Blair (2006) concludes that the prefrontal cortex, where fluid abilities are located, declines more rapidly than other brain areas. Reasons for a general cognitive decline are seen in the rate of information processing and processing speed (Baltes, Lindenberger, & Staudinger, 2006; Salthouse, 1996), which can explain changes in fluid reasoning (Zimprich & Martin, 2002). Tucker-Drob (2011) identified global but also specific changes for domains such as reasoning and processing speed.

3.2.2. Learning opportunities

Empirical studies show that schooling has impact on acquired skills such as reading skills (Crone & Whitehurst, 1999; Gustafsson, 2016; Rasmusson, Albæk, Lind, & Myrberg, 2018; Sulkunen & Malin, 2018), while differences in schooling are also related to differences in intelligence (Becker, Lüdtke, Trautwein, Köller, & Baumert, 2012; Ceci & Williams, 1997; Guill, Lüdtke, & Köller, 2017). These effects also remain in adulthood (Clouston et al., 2012; Ritchie, Bates, Der, Starr, & Deary, 2013). Effects from schooling on both – intelligence measures such as fluid reasoning, but also on measures that contain also acquired skills classically referred to as crystallized skills – are not surprising when the relationship of fluid abilities, such as reasoning, and crystallized abilities, such as acquired knowledge, is considered. Reasoning explains high amounts of variance of later educational achievement (Deary, Strand, Smith, & Fernandes, 2007). According to investment theory (Cattell, 1987), fluid intelligence describes a "general power" (p. 138) that influences crystallized skill levels in the future together with personality and motivation (Cattell, 1963). Schweizer and Koch (2002) emphasize that this relation is established through learning. Although Ferrer and McArdle (2004) conclude that the relationship of fluid and crystallized skills is much more complex than initially thought, they support the importance of fluid skills for achievement. If effects of education on different skills are compared, effects of education were rather on crystallized knowledge than on fluid skills (Finn et al., 2014) and rather on knowledge than on elementary processes (Ritchie et al., 2013). Also a meta-analysis (Ritchie & Tucker-Drob, 2018) supports, that education is related to the performance in achievement tests (directly related to what is taught at school) and typical IQ tests. If effects were compared, they were rather larger on the group of achievement tests than on the group of IQ-tests.

4. Hypotheses

The first step of the present study is to investigate the dimensionality of PIAAC competencies when analyzed together with reasoning and processing speed. If domain-specific skills influence test results of the PIAAC competence tests, a model with additional nested factors should describe the data better than a single factor. Since PIAAC competencies are assumed to be learnable and to be based also on knowledge, it is expected that *test scores from PIAAC literacy and numeracy represent beyond what they have in common with reasoning and processing speed also individual differences in domain-specific skills (Hypothesis 1)*.

Given results support this separability, a common factor of PIAAC competencies with reasoning and processing speed should be interpreted as representing rather general aspects of cognitive skills that are close to what is classically referred to as fluid abilities. The nested, domain-specific factors, in turn, are assumed to represent rather domain-specific skills and to be close to what is classically referred to as crystallized abilities. The validity of this intended interpretation of the nested factors is investigated by analyzing relations to other variables (AERA, APA, & NCME, 2014).

Since cognitive decline seems to be different for fluid and crystallized abilities, different relations to age are expected for the factors separated in the dimensionality analyses. The common factor is assumed to capture global cognitive decline, as it represents common aspects between PIAAC competencies with reasoning and processing speed. Domain-specific decline should be captured by the nested factors. Based on predictions from theory and empirical findings *we expect for PIAAC competencies that - beyond the relation of the common factor to age - the domain-specific factors are unrelated to age (Hypothesis 2)*. This would support that PIAAC competencies assess beyond their overlaps with general cognitive skills rather crystallized aspects and strengthens the intended test score interpretation of test scores from PIAAC competencies as assessing domain-specific knowledge.

In order to investigate whether PIAAC competencies can be interpreted as being learned and as representing domain-specific aspects, *we expect for PIAAC competencies that - beyond the relation of the common factor to learning opportunities during one's lifetime - the domain-specific factors are positively related to learning opportunities during one's lifetime (Hypotheses 3)*. If this was not the case, PIAAC competencies would not be affected by learning opportunities above the general effect on general cognitive skills, the common basis of all tests. If competencies contain domain-specific acquired skills, they should be related to learning opportunities.

The domain-specificity of competencies assessed in large-scale assessment studies is questioned by large correlations between domains (at the country-level) and between ratings on the content of the domains (Rindermann, 2006; Rindermann & Baumeister, 2015). Hence we also address the question of whether the nested factors represent domain-specific aspects. Persons who perform competence tasks in daily life more frequently might be better at performing those tasks (Hartshorne & Germine, 2015; Nisbett et al., 2012). However, also the other direction is possible that persons who are better in certain competencies tend to select those tasks in everyday life more frequently. Although no direction should be taken on in this study, how nested factors relate to domain-specific variables should be considered to describe whether the nested factors describe domain-specific aspects. *We expect for PIAAC competencies that - beyond the relation of the common factor to the use of competence specific skills in daily life - the domain-specific factors are positively related to the use of competence specific skills in daily life (Hypotheses 4)*.

5. Method

5.1. Sample and procedures

A subsample of German PIAAC 2012 participants completed in 2015 the computer-based PIAAC literacy and numeracy tests (German PIAAC longitudinal study; GESIS – Leibniz Institute for the Social Sciences, German Socio-Economic Panel (SOEP) at DIW Berlin and LIfBi – Leibniz Institute for Educational Trajectories (2017); Rammstedt, Martin, Zabal, Carstensen, & Schupp, 2017), and in 2016 a test for perceptual speed (SOEP Symbol-Digit-Test; Schupp, Herrmann, Jaensch, & Lang, 2008) and a reasoning test (number series test; Engelhardt & Goldhammer, 2018). Data from $N = 903$ respondents who completed one or both of the PIAAC competence assessments (random assignment), the Symbol-Digit-Test, and the number series test, were available for analyses. From those participants, $n = 447$ were male and $n = 456$ were female. Further information concerning the design of the PIAAC-L study can be found in Rammstedt et al. (2017).

5.2. Measures

5.2.1. Performance tests

PIAAC literacy tasks[1] contain continuous and non-continuous texts and PIAAC numeracy tasks[2] require responding to mathematical problems in real contexts (OECD, 2016a, p. 18). Each test included 49 items, which were administered in a two-stage adaptive test design (Kirsch & Yamamoto, 2013, pp. 10) so that each respondent worked on 20 items per competence domain (Zabal, Martin, & Rammstedt, 2017). Three variables, education level, being a native speaker, and the passing score on the computer-based assessment core tasks (cf. OECD, 2016a), increased the probability that the respondent received a testlet appropriate to their skill level and were, thus, included as auxiliary variables (Enders, 2010) into the data analyses. Items were dichotomously scored. Fitting 2-parameter IRT models in Mplus (Muthén & Muthén, 2015) revealed that all literacy items ($M = .62$, $SD = .10$) and all numeracy items ($M = .59$, $SD = .12$) each loaded significantly on a latent ability

[1] For example items, see www.oecd.org/skills/piaac/Literacy%20Sample%20Items.pdf (accessed February 13, 2021).

[2] For example items, see www.oecd.org/skills/piaac/Numeracy%20Sample%20Items.pdf (accessed February 13, 2021).

factor.

To represent fluid reasoning, a number series test (McArdle & Woodcock, 2009) was used. All of the 15 dichotomously scored items loaded significantly on a latent ability factor (*M* = .69, *SD* = .11). For perceptual speed, respondents had to recode symbols (Symbol-Digit test; Schupp et al., 2008) and the sum of correct digits within three time intervals was treated as indicators for the analyses. All three indicators loaded significantly on a latent factor (*M* = .86, *SD* = .05).

5.2.2. Background variables

Variables that were included in the data analyses for Hypotheses 2–4 are described in the following. The age of the participants in the assessment in 2015 was between 19 and 69 years (*M* = 42.57, *SD* = 13.61). To capture learning opportunities in both formal and informal settings, three different aspects should be covered: For formal settings, these are the quality of education and the length of attending educational settings, as they seem to impact adult literacy and numeracy performance (Gustafsson, 2016; Sulkunen & Malin, 2018). Quality of education was an eight-level variable which describes the highest obtained level of formal education, from primary, lower secondary, upper secondary, post-secondary (non-tertiary) degrees and four different tertiary degrees, which were professional, bachelor, master and research degree (*M* = 4.26, *SD* = 1.74). To describe the length of schooling, the number of years that the respondents have spent in educational settings, school or professional, was used (*M* = 14.54, *SD* = 3.10). To also capture informal settings, the number of books at home when the respondents were 16 years old was used (6 categories; *M* = 3.33, *SD* = 1.41). To assess skill use, respondents were asked on a scale from 1 to 5 how often they visit libraries, book stores, or look for books online (*M* = 3.71, *SD* = 1.02) and whether they use their mathematical skills or need to deal with numbers during work (yes: *n* = 631, no: *n* = 101).

5.2.3. Data analyses

Data was analyzed using Mplus (Muthén & Muthén, 2015). Structural equation models with categorical items were applied using the MLR estimator. Three variables were responsible for adaptive testlet selection (education level, native speaker, computer-based assessment core score). Disregarding these variables when modelling the data would mean that missingness depends on unobserved data violating the missing at random (MAR) assumption. Because more able respondents tend to receive harder items and less able respondents receive easier items, theta estimates would be underestimated for able and overestimated for less able respondents. To account for that, we included those variables that were responsible for testlet selection as auxiliary variables into the analyses, making the assumption justifiable that the not-administered items were MAR. In this case, the MLR estimator is adequate to handle missing data that resulted from the adaptive test design of the PIAAC competence tests. According to Enders (2010), correlations should be preferred to predictors, as they do not change the interpretation of the estimates. Correlations of the three variables for testlet selection with the latent factors, and among each other, were included in all models.

For investigating Hypothesis 1, various models were tested and compared. First, a common factor for all four performance tests was modeled. For modeling nested factors, bifactor-(S – 1) models were used as they are advantageous regarding anomalous results, such as high standard errors leading to invalid conclusions, vanishing specific factors, or irregular loading patterns (Eid, Geiser, Koch, & Heene, 2017; Eid, Krumm, Koch, & Schulze, 2018). In those models, nested factors for all domains but the reference domain are modeled, which can be chosen based on theoretical assumptions and goals. In the present study, fluid reasoning was chosen as reference domain because it is assumed to represent best what all tests have in common. Second, we modeled nested factors for literacy, numeracy, and perceptual speed, and tested, third, a more parsimonious model, which did not contain a separate nested factor for perceptual speed. A four-dimensional model with four latent factors for all four performance tests (literacy, numeracy, reasoning, and perceptual speed) serves as a reference for the interpretation of loadings. In this model, all latent factors were allowed to correlate with each other.

Explanatory analyses (Hypotheses 2–4) are based on the best fitting model from the model comparisons. Age was included in all models as a predictor, as it is likely that cohort effects, such as different learning opportunities for different cohorts, influence the effects of opportunities to learn. While age and variables for learning opportunities were included as predictors (Hypotheses 2 and 3), skill use variables (Hypotheses 4) were included as correlating variables, because no effect in one direction was assumed between variables describing the skill use and competence in the respective domain. For the skills use variables, the polarity of effects was inverted so that positive correlations indicate that a more frequent use is related to higher scores on the latent factors. For a better interpretation of the results, we show also the relation of all background variables from Hypotheses 2–4 to the four latent factors from the four-dimensional model (cf. Table 4).

6. Results

6.1. Dimensionality analysis (Hypothesis 1)

To approach the dimensionality analyses, four CFA models were fitted. The results of all models are presented in Table 1. Fig. 1 shows the nested factor model with three nested factors.

According to the information criteria (AIC, BIC, and the sample-size adjusted BIC), the model with three nested factors fitted the data best supporting that PIAAC competence items represent common aspects beyond their overlaps with the reasoning and perceptual speed test (Hypothesis 1). The literacy (σ^2 *(SE)* = 2.88 (1.00), *p* = .004) but not the numeracy factor (σ^2 *(SE)* = 5.61 (3.65), *p* = .124), explained the shared variance between the items beyond the common factor. Also the common factor (σ^2 *(SE)* = 4.39 (1.08), *p* < .001) and perceptual speed factor (σ^2 *(SE)* = 0.54 (0.04), *p* < .001) explained significant variance components. A couple of items did not load significantly on their nested factors (4 for literacy and 21 for numeracy). In the four-dimensional model, all items loaded on their specified factors (cf. Table 2) and all latent factors represented significant variance components.

The analyses for Hypotheses 2 through 4 are intended to provide explanations for their separability and interpretations for the latent factors beyond this statistical distinction, which would also rule out an alternative explanation of modeling method factors. Although the numeracy factor did not capture significant proportions of variance between the items beyond the common factor, analyses are still conducted, but limitations regarding the interpretation of this factor are considered in the discussion.

6.2. Relations to age (Hypothesis 2)

To address Hypothesis 2, all latent factors were regressed on age (cf. Table 3). As expected (Hypothesis 2), age was not predictive for the nested competence factors for literacy (β = −0.08, *p* = .400) and also not for numeracy (β = 0.08, *p* = .576). This supports the assumption that the shared variance components of PIAAC literacy und numeracy items that go beyond overlaps with reasoning and perceptual speed show a pattern that is typically expected for crystallized aspects. Instead, age was negatively predictive for the common factor (β = −0.24, *p* < .001) and the nested factor of perceptual speed (β = −0.60, *p* < .001). Within the four-dimensional model, all factors were negatively related to age (cf. Table 4).

6.3. Relations to formal and informal learning opportunities (Hypotheses 3)

To address Hypotheses 3, the predictive value of formal and informal

Table 1
Hypothesis 1 – dimensionality analyses.

Model	Latent Factors	Information Criteria	Variance (SE), sign.	Items from Domain…	n.s. Loadings[1]	Descriptives of standardized Loadings			
						M	*SD*	*Min*	*Max*
1) Common Factor Model	Common	AIC 50503	4.48 (0.93), $p < .001$	Literacy	0/49	.58	.10	.31	.78
		BIC 51690		Numeracy	0/49	.57	.12	.30	.77
		BIC (adj.) 50905		Reasoning	0/15	.59	.10	.40	.78
				Perceptual Speed	0/3	.44	.01	.43	.45
2) Three Nested Factors	Common	AIC 48884	4.39 (1.08), $p < .001$	Literacy	0/49	.49	.10	.28	.71
		BIC 50600		Numeracy	4/49	.49	.13	.18	.71
		BIC (adj.) 49466		Reasoning	0/15	.66	.11	.46	.90
				Perceptual Speed	0/3	.40	.01	.38	.41
	Literacy		2.88 (1.00), $p = .004$		4/49	.39	.14	.13	.83
	Numeracy		5.61 (3.65), $p = .124$		21/49	.32	.12	.10	.63
	Percept. Sp.		0.54 (0.04), $p < .001$		0/3	.76	.06	.70	.80
3) Two Nested Factors	Common	AIC 50141	4.33 (1.05), $p < .001$	Literacy	0/49	.49	.09	.29	.69
		BIC 51828		Numeracy	4/49	.48	.12	.22	.69
		BIC (adj.) 50713		Reasoning	0/15	.66	.11	.45	.92
				Perceptual Speed	0/3	.53	.01	.52	.54
	Literacy		2.88 (0.97), $p = .003$		2/49	.39	.14	.16	.69
	Numeracy		4.51 (2.40), $p = .060$		8/49	.35	.12	.10	.62

Note: [1] p > = .05.

learning opportunities was analyzed (cf. Table 3). As expected, all three variables were positively predictive for the nested factor of literacy (educational attainment: $\beta = 0.33$, $p < .001$; number of years in educational settings: $\beta = 0.24$, $p = .008$; number of books: $\beta = 0.30$, $p < .001$) but against expectation not the nested factor for numeracy (educational attainment: $\beta = 0.16$, $p = .130$; number of years in educational settings: $\beta = 0.07$, $p = .602$; number of books: $\beta = 0.13$, $p = .129$), supporting that at least PIAAC literacy items are also above commonalities with reasoning, perceptual speed, and numeracy, related to learning opportunities. The educational variables were also positively predictive for the common factor but not predictive for the nested perceptual speed factor (cf. Table 3). Also in the four-dimensional model, all factors were positively related to the three variables assessing learning opportunities, except for one not significant relation of years in education and perceptual speed (cf. Table 4).

6.4. Relations to domain-specific variables (Hypotheses 4)

As expected (Hypothesis 4a), persons who visited libraries or search online for books more often had higher scores (cf. Table 5) on the nested

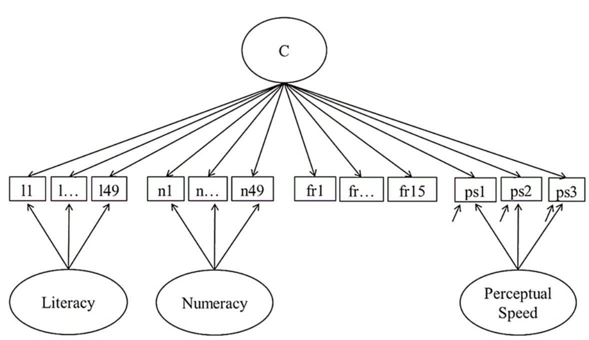

Fig. 1. Nested factor model. A common factor with additional nested factors for literacy, numeracy, and perceptual speed.

Table 2
Four-dimensional model.

Model	Latent Factors	Variance (SE), sign.	n.s. Loadings[1]	Descriptives of standardized Loadings			
				M	*SD*	*Min*	*Max*
4 Factors	Literacy	6.36 (1.51), $p < .001$	0/49	.61	.10	.33	.81
	Numeracy	4.51 (1.41), $p = .001$	0/49	.59	.12	.33	.80
	Reasoning	8.91 (1.85), $p < .001$	0/15	.69	.11	.50	.89
	Percept. Sp.	0.66 (0.04), $p < .001$	0/3	.86	.04	.81	.89

Note: [1] *p* > = .05.

Table 3
Regressions of latent factors based on the Model with Three Nested Factors (Hypotheses 2-3).

		Common Factor				Literacy				Numeracy				Perceptual Speed			
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
Hyp. 2	Age	**−0.24**	.05	−4.50	< .001	**−0.08**	.09	−0.84	.400	**0.08**	.15	0.56	.576	**−0.60**	.03	−21.73	< .001
	R^2		.06*				.01 n.s.				.01 n.s.				.36***		
Hyp. 3a	Educational Attainment	**0.39**	.04	9.36	< .001	**0.33**	.07	4.72	<.001	**0.16**	.11	1.51	.130	**0.03**	.04	0.86	.390
	Age	−0.24	.05	−4.35	< .001	−0.13	.08	−1.56	.120	0.02	.13	0.15	.880	−0.61	.03	−21.26	< .001
	R^2		.18***				.11*				.03 n.s.				.37***		
Hyp. 3b	Years in Education	**0.15**	.05	2.99	.003	**0.24**	.09	2.65	.008	**0.07**	.13	0.52	.602	**0.04**	.04	0.81	.420
	Age	−0.22	.05	−4.05	< .001	−0.07	.09	−0.72	.470	0.06	.13	0.45	.651	−0.60	.03	−21.86	< .001
	R^2		.07*				.06 n.s.				.01 n.s.				.36***		
Hyp. 3c	Books at Age 16	**0.17**	.04	3.98	< .001	**0.30**	.07	4.57	< .001	**0.13**	.08	1.52	.129	**0.05**	.04	1.30	.193
	Age	−0.21	.05	−4.08	< .001	−0.08	.08	−0.94	.345	0.08	.13	0.59	.558	−0.60	.03	−22.30	< .001
	R^2		.08**				.10*				.02 n.s.				.37***		

Note: $p < .001$***; $p < .01$**; $p < .05$*; R^2 = proportion of explained variance in the latent factor by background variables.

Table 4
Correlations of latent factors of a four-dimensional model and background variables.

	Literacy	Numeracy	Fluid Reasoning	Perceptual Speed
Literacy	5.24 (1.19)***	–	–	–
Numeracy	.88***	4.10 (1.27)**	–	–
Fluid Reasoning	.66***	.74***	8.17 (1.71)***	–
Perceptual Speed	.43***	.39***	.38***	0.66 (0.04)***
Age	−.21***	−.15***	−.15***	−.61***
Educat. Attainment	.31***	.32***	.24***	.11**
Years in Education	.18***	.12*	.13**	.05 n.s.
Books at Age 16	.27***	.18***	.16***	.08**
Literacy in Leisure[1]	.14***	.06 n.s.	.05 n.s.	.04 n.s.
Numeracy at Work[1]	.11**	.16***	.08 n.s.	.04 n.s.
n.s. loadings ($p >$ .05)	0/49	0/49	0/15	0/3

Note: $p < .001$***; $p < .01$**; $p < .05$*. The diagonal contains variances of the latent factors. [1]A more frequent use was coded in the data set with lower values. We inverted the polarity of the *r*-values for these two variables in the table. Positive correlations describe that a more frequent use is related to higher test scores.

factor for literacy ($r = 0.24$, $p < .001$), but not for numeracy ($r = -0.04$, $p = .589$) and persons who use numeracy skills at work more often (Hypothesis 4b), had higher values on the nested numeracy factor ($r = 0.19$, $p = .003$) but not on the nested literacy factor ($r = -0.06$, $p = .343$). This supports that PIAAC literacy and numeracy items assess beyond commonalities with all the other items aspects that are specific to the domain. Both skill use variables were also related to the common factor (literacy in leisure: $r = 0.10$, $p = .007$; numeracy at work: $r = 0.09$, $p = .021$), but not to the nested perceptual speed factor (literacy in leisure: $r = 0.03$, $p = .472$; numeracy at work: $r = -0.01$, $p = .836$). In the four-dimensional model (cf. Table 4), the domain-specific variables were only related to their specific competence factor, while the use of numeracy at work was also related to the literacy factor.

7. Discussion

7.1. Test score interpretation of PIAAC competencies

The focus of the present study was to collect validity evidence for the intended test score interpretation of PIAAC competencies as being learnable and to be based on knowledge (OECD, 2016a). Since competencies are characterized by complexity and contextualized items ask for general but also specific skills (Klieme et al., 2008; OECD, 2016a), an approach was chosen which divides PIAAC competencies always in an overlapping part with general cognitive skills, and aspects beyond those overlaps. Predictions of the present study were based exactly on aspects beyond those overlaps, because they describe whether PIAAC competencies differ from general cognitive skills (focusing on the internal structure) and allow describing those specific aspects in greater detail (focusing on relations to other variables).

At first, a nested factor model fitted the data better than a single factor model supporting that the PIAAC competencies literacy and numeracy assessed common aspects beyond overlaps with reasoning and perceptual speed, collecting validity evidence based on the internal structure (AERA, APA, & NCME, 2014). At second, whether the nested factors fulfilled expected relations with other variables based on predictions from theory and empirical findings was investigated. For the nested factor for PIAAC literacy, the result pattern was as expected. Being unrelated to age, positively related to informal and formal learning opportunities and positively related to skill use, the nested factor for PIAAC literacy fulfilled all expectations for a factor that should be interpreted as learnable and to be based on acquired knowledge. The nested factor for PIAAC numeracy fulfilled theses expectations not entirely: although being as expected unrelated to age and positively related to skill use, the nested numeracy factor was against expectation unrelated to learning opportunities. Considered together with results from dimensionality analyses – the nested factor for PIAAC numeracy did not represent significant variance components and some numeracy items did not load on the factor – these findings suggest that numeracy items may share rather little beyond overlaps with PIAAC literacy, reasoning, and perceptual speed and that these shared aspects did not fulfill all expectations for a factor that should be interpreted as learnable and to be based on acquired knowledge. Therefore, results support validity evidence for the intended test score interpretation based on relations to other variables (AERA, APA, & NCME, 2014) for PIAAC literacy but not entirely for PIAAC numeracy.

These conclusions are also supported by results for the common factor and the nested factor for perceptual speed. The separation into

Table 5
Correlations of latent factors based on the Model with Three Nested Factors (Hypotheses 4).

		Common Factor				Literacy				Numeracy				Perceptual Speed			
		r	*SE*	*z*	*p*	*r*	*SE*	*z*	*p*	*r*	*SE*	*z*	*p*	*r*	*SE*	*z*	*p*
Hyp. 4a	Literacy in Leisure	**0.10**[1]	.04	2.71[1]	.007	**0.24**[1]	.07	3.75[1]	<.001	**−0.04**[1]	.07	−0.54[1]	.589	**0.03**[1]	.04	0.72[1]	.472
	Age	−0.24	.05	−4.41	< .001	−0.08	.09	−0.83	.407	0.07	.15	0.48	.628	−0.60	.03	−21.79	<.001
Hyp. 4b	Numeracy at Work	**0.09**[1]	.04	2.32[1]	.021	**−0.06**[1]	.06	−0.95[1]	.343	**0.19**[1]	.07	2.99[1]	.003	**−0.01**[1]	.04	−0.21[1]	.836
	Age	−0.24	.05	−4.66	<.001	−0.08	.09	−0.84	.400	0.07	.14	.51	.613	−0.60	.03	−21.79	<.001

Note: [1]A more frequent use was coded in the data set with lower values. We inverted the polarity of the *r*- and *z*-values in the table. Positive correlations describe that a more frequent use is related to higher test scores.

general cognitive skills and specific skills through dimensionality analysis and also the predictions for relations to other variables were based on the assumption that a common factor captures a common basis of all skills and is close to fluid reasoning. If this is the case, the remaining commonalities between PIAAC literacy and numeracy items beyond the common factor can be then interpreted as domain-specific aspects distinguished from general cognitive skills. Results support this interpretation of the common factor as being close to reasoning: At first, because all items had significant loadings on this common factor in the nested factor model, except for four numeracy items. Hence, this factor represents skills required by all tests. Besides, this factor was negatively related to age, which is in line with expectations from literature expected for fluid skills (Horn & Cattell, 1967; Li et al., 2004; McArdle et al., 2002; Verhaeghen & Salthouse, 1997). The common factor was also positively related to all three educational variables, which is in line with literature supporting that formal and informal learning opportunities are related to a positive cognitive development (Ceci & Williams, 1997; Guill et al., 2017; Ritchie & Tucker-Drob, 2018). Although gains in fluid reasoning seem to be best reached when working memory is trained, effects of schooling may also act via training of self-regulation skills (Nisbett et al., 2012).

Also the nested factor for perceptual speed fulfilled expectations based on predictions from literature and showed compared to the nested factors of the PIAAC competencies a completely different pattern: The nested factor for perceptual speed was above the common factor negatively related to age, and the perceptual speed factor showed also in the four-dimensional model the strongest negative relation to age, which is in line with previous empirical findings (Lang et al., 2007; Salthouse, 1996; Zimprich & Martin, 2002). Also Verhaeghen and Salthouse (1997) concluded based on their meta-analyses, that largest age-related effects were found for speed of processing, more than for reasoning. The different relations to age for the nested factors are also in line with empirical findings on different growth patterns for processing speed, fluid, and crystallized skills (Li et al., 2004; McArdle et al., 2002). Also the not significant relation of the nested factor for perceptual speed to variables describing learning opportunities is in line with findings from literature (Lang et al., 2007). Effects of education were rather found on skills such as reasoning or knowledge than on elementary cognitive processes (Ritchie, Bates, & Deary, 2015; Ritchie et al., 2013). Interestingly, the perceptual speed factor in the four-dimensional model was indeed positively related to two out of three variables describing learning opportunities. This supports that the chosen approach – to disentangle general aspects in cognitive tests from specific aspects – leads to a more differentiated and clearer pattern when relations to other variables are considered.

Taken together, the intended interpretation of test scores from PIAAC literacy is supported by results referring to the relation of the nested factor of PIAAC literacy to other variables and is also supported by the result pattern of the other factors.

7.2. Contribution to the field

Results from the present study contribute at first to the question how PIAAC competencies can be interpreted. Given the great amount of possible research questions that can be answered using large-scale assessment studies such as PIAAC, validity questions of the competence tests are of high importance. Second, studies on validity issues focused by now mostly on student assessments such as PISA and little is known about the validity of the test score interpretation of large-scale assessments on adults. Focusing on the adult population allows extending approaches used for student data (e.g. dimensionality analyses), because significant variables from the lifespan can be related to the separated factors when the adult population is addressed. Still, it has to be kept in mind that results from the adult population cannot be necessarily transferred to students. Third, the present study also contributes to the discussion about the closeness of competencies assessed in large-scale assessment studies and intelligence (cf. Rindermann, 2006; Rindermann & Baumeister, 2015). For test scores from PIAAC literacy it can be concluded that they assess, beyond overlaps with skills that might be referred to as intelligence, aspects that show a pattern typically expected for learned skills and acquired knowledge. Fourth, the used approach of relating nested factors to other variables has proven to be helpful to reach a differentiated pattern of relations to other variables for the common and nested factors and is considered to be advantageous compared to conclusions that could be drawn only based on a multi-dimensional factor model. Such an approach could be also useful for validating test score interpretations from other complex skills, such as twenty first century skills (Binkley et al., 2012).

7.3. Limitations and future research

As this study was conducted within the frame of PIAAC-L in Germany (cf. Rammstedt et al., 2017), results supporting the construct interpretations of the PIAAC competencies are, at first, only valid to German samples. A potential disadvantage stems from the time delay of one year between the competence and the general cognitive skills assessment. However, we assume that this does not question the validity of results because one year is rather short time within the entire lifespan for an adult sample. With data from the PIAAC study we rely also on a cross sectional data set, which means that negative relationships to age cannot be interpreted causally as cognitive decline in our data set.

The conclusions that can be drawn regarding the validity of the test score interpretation are only clear for PIAAC literacy, but not for PIAAC numeracy. Further research may focus closer on the validity of test scores from PIAAC numeracy, which were in the present study not clearly distinguishable from what items from PIAAC literacy, reasoning, and perceptual speed had in common. For complex skills such as PIAAC competencies, also experimental approaches (cf. Engelhardt, Goldhammer, Naumann, & Frey, 2017) could be applied to address further sources of validity evidence, such of response processes (AERA, APA, & NCME, 2014). Since different abilities decline at different ages (Hartshorne & Germine, 2015) and the PIAAC sample describes a large age range, further validation efforts might also focus on the question whether the conclusions hold for different age groups.

8. Conclusion

Taken together, this study provided validity evidence for the intended test score interpretation of PIAAC competencies - as being learnable and based on acquired knowledge - based on the internal structure and relations to other variables for PIAAC literacy, while results for PIAAC numeracy were less clear. The chosen approach of separating general and specific aspects in PIAAC competencies first, and relating them second to variables from the lifespan for which clear expectations exist, seemed to be useful to meet the challenge of addressing validity issues in complex constructs such as competencies. Herewith this study contributes to the question of overlaps between competencies and general cognitive skills by providing results based on an adult data set from the PIAAC study.

Funding sources

The Number Series Study was funded by the Leibniz Association (Grant Number SAW-2015-GESIS-2) and the Centre for International Student Assessment (ZIB; Grant Number ZIB2016).

References

American Educational Research Association, American Psychological Association & National Council on Measurement in Education [AERA, APA, & NCME]. (2014). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Life span theory in developmental psychology. In W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology, theoretical models of human development: Vol. 1. Theoretical models of human development* (pp. 569–664). Hoboken, NJ: John Wiley & Sons, 6. Aufl., S.

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*, 165–176. https://doi.org/10.1016/j.edurev.2009.04.002.

Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology, 104*, 682–699. https://doi.org/10.1037/a0027608.

Binkley, M., Erstad, O., Herman, J., Raizen, R., Ripley, M., & Rumble, M. (2012). Defining 21st century skills. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 1st century skills* (pp. 17–66). Dordrecht: Springer.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *The Behavioral and Brain Sciences, 29*, 109–160. https://doi.org/10.1017/S0140525X06009034.

Brunner, M. (2005). *Mathematische Schülerleistung: Struktur, Schulformunterschiede und Validität [Mathematics achievement: Structure, differences between academic tracks, and validity]*. Doctoral Dissertation. Berlin, Germany: Humboldt University.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1–22. https://doi.org/10.1037/h0046743.

Cattell, R. B. (1987). *Intelligence: Its structure, growth and action*. Amsterdam: North-Holland.

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist, 52*, 1051–1058. https://doi.org/10.1037/0003-066X.52.10.1051.

Clouston, S. A., Kuh, D., Herd, P., Elliott, J., Richards, M., & Hofer, S. M. (2012). Benefits of educational attainment on adult fluid cognition: International evidence from three birth cohorts. *International Journal of Epidemiology, 41*, 1729–1736. https://doi.org/10.1093/ije/dys148.

Crone, D. A., & Whitehurst, G. J. (1999). Age and schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology, 91*(4), 604–614.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13–21. https://doi.org/10.1016/j.intell.2006.02.001.

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in g-factor models: Explanations and alternatives. *Psychological Methods, 22*, 541–562. https://doi.org/10.1037/met0000083.

Eid, M., Krumm, S., Koch, T., & Schulze, J. (2018). Bifactor models for predicting criteria by general and specific factors: Problems of nonidentifiability and alternative solutions. *Journal of Intelligence, 6*, 42. https://doi.org/10.3390/jintelligence6030042.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford press.

Engelhardt, L., & Goldhammer, F. (2018). *Number series study (DIPF): Technical report*. GESIS Papers 2018|01. Cologne: GESIS.

Engelhardt, L., Goldhammer, F., Naumann, J., & Frey, A. (2017). Experimental validation strategies for heterogeneous computer-based assessment items. *Computers in Human Behavior, 76*, 683–692. https://doi.org/10.1016/j.chb.2017.02.020.

Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology, 40*, 935–952. https://doi.org/10.1037/0012-1649.40.6.935.

Finn, A. S., Kraft, M. A., West, M. R., Leonard, J. A., Bish, C. E., Martin, R. E., … Gabrieli, J. D. (2014). Cognitive skills, student achievement tests, and schools. *Psychological Science, 25*, 736–744. https://doi.org/10.1177/0956797613516008.

GESIS – Leibniz Institute for the Social Sciences, German Socio-Economic Panel (SOEP) at DIW Berlin & LIfBi – Leibniz Institute for Educational Trajectories. (2017). *PIAAC-Longitudinal (PIAAC-L), Germany. GESIS Data Archive, Cologne. ZA5989 Data file Version 3.0.0*. https://doi.org/10.4232/1.12925.

Guill, K., Lüdtke, O., & Köller, O. (2017). Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study. *Learning and Instruction, 47*, 43–52. https://doi.org/10.1016/j.learninstruc.2016.10.001.

Gustafsson, J. E. (2016). Lasting effects of quality of schooling: Evidence from PISA and PIAAC. *Intelligence, 57*, 66–72. https://doi.org/10.1016/j.intell.2016.05.004.

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science, 26*, 433–443. https://doi.org/10.1177/0956797614567339.

Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J.-A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research, 62*, 333–353. https://doi.org/10.1080/00313831.2016.1258726.

Horn, J. L. (1965). *Fluid and crystallized intelligence: A factor analytic and developmental study of the structure among primary mental abilities*. Unpublished doctoral dissertation. Champaign: University of Illinois.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57*, 253–270. https://doi.org/10.1037/h0023816.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107–129. https://doi.org/10.1016/0001-6918(67)90011-X.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Kirsch, E., & Yamamoto, K. (2013). Chapter 1: PIAAC assessment design. *OECD Technical report of the survey of adult skills (PIAAC)*. Retrieved from OECD website: https://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf.

Klieme, E., & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik, 52*(6), 876–903.

Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 3–22). Hogrefe Publishing.

Lang, F. R., Weiss, D., Stocker, A., & von Rosenbladt, B. (2007). Assessing cognitive capacities in computer-assisted survey research: Two ultra-short tests of intellectual ability in the German Socio-Economic Panel (SOEP). *Schmollers Jahrbuch, 127*(1), 183–192.

Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15*, 155–163. https://doi.org/10.1111/j.0956-7976.2004.01503003.x.

McArdle, J. J., & Woodcock, R. W. (2009). *Adaptive testing with selected Woodcock cognitive tests [Powerpoint slides]*. Charlottesville, VA: Longitudinal Research Institute.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38*, 115. https://doi.org/10.1037//0012-1649.38.1.115.

McArdle, J. J., Hamagami, F., Meredith, W., & Bradway, K. P. (2000). Modeling the dynamic hypotheses of Gf–Gc theory using longitudinal life-span data. *Learning and Individual Differences, 12*, 53–79. https://doi.org/10.1016/S1041-6080(00)00036-4.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10. https://doi.org/10.1016/j.intell.2008.08.004.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus user's guide* (seventh edition). Los Angeles, CA: Muthén & Muthén.

Nagy, G. (2006). *Berufliche Interessen, kognitive und fachgebundene Kompetenzen: Ihre Bedeutung für die Studienfachwahl und die Bewährung im Studium [Vocational interests, cognitive and scholastic abilities: Their role in choice of major and success at university]*. Unpublished doctoral dissertation. Berlin, Germany: Free University.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., … Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist, 67*, 130–159. https://doi.org/10.1037/a0026699.

OECD. (2016a). *The survey of adult skills: Reader's companion* (second edition). Paris: OECD Skills Studies, OECD Publishing. https://doi.org/10.1787/9789264258075-en.

OECD. (2016b). *PISA 2015 results (Volume I): Excellence and equity in education*. Paris: PISA, OECD Publishing. https://doi.org/10.1787/9789264266490-en.

Prenzel, M., Walter, O., & Frey, A. (2007). PISA misst Kompetenzen [PISA measures Competences]. *Psychologische Rundschau, 58*, 128–136. https://doi.org/10.1026/0033-3042.58.2.128.

Rammstedt, B., Martin, S., Zabal, A., Carstensen, C., & Schupp, J. (2017). The PIAAC longitudinal study in Germany: Rationale and design. *Large-scale Assessments in Education, 5*, 1–11. https://doi.org/10.1186/s40536-017-0040-z.

Rasmusson, M., Albæk, K., Lind, P., & Myrberg, M. (2018). Cognitive foundation skills following vocational versus general upper-secondary education: A long-term perspective. *Scandinavian Journal of Educational Research*, 1–22. https://doi.org/10.1080/00313831.2018.1466361.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*, 1–48. https://doi.org/10.1006/cogp.1999.0735.

Rindermann, H. (2006). Was messen internationale Schulleistungsstudien? Schulleistungen, Schülerfähigkeiten, kognitive Fähigkeiten, Wissen oder allgemeine Intelligenz? [What do international academic achievement studies measure? Academic achievement, student abilities, cognitive abilities, knowledge, or general intelligence?]. *Psychologische Rundschau, 57*, 69–86. https://doi.org/10.1026/0033-3042.57.2.69.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21*, 667–706. https://doi.org/10.1002/per.634.

Rindermann, H., & Baumeister, A. E. E. (2015). Validating the interpretations of PISA and TIMSS tasks: A rating study. *International Journal of Testing, 15*, 1–22. https://doi.org/10.1080/15305058.2014.966911.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science, 29*, 1358–1369. https://doi.org/10.1177/0956797618774253.

Ritchie, S. J., Bates, T. C., & Deary, I. J. (2015). Is education associated with improvements in general cognitive ability, or in specific skills? *Developmental Psychology, 51*, 573–582. https://doi.org/10.1037/a0038981.

Ritchie, S., Bates, T., Der, G., Starr, J., & Deary, I. (2013). Education is associated with higher later life IQ scores, but not with faster cognitive processing speed. *Psychology and Aging, 28*, 515–521. https://doi.org/10.1037/a0030820.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403–428. https://doi.org/10.1037/0033-295X.103.3.403.

Saß, S., Kampa, N., & Köller, O. (2017). The interplay of g and mathematical abilities in large-scale assessments across grades. *Intelligence, 63*, 33–44. https://doi.org/10.1016/j.intell.2017.05.001.

Schupp, J., Herrmann, S., Jaensch, P., & Lang, F. R. (2008). Erfassung kognitiver Leistungspotentiale Erwachsener im Sozio-oekonomischen Panel (SOEP) [Surveying Cognitive Achievement Potentials of Adults in the Socio-Economic Panel (SOEP)]. *Data Documentation* (p. 32). DIW Berlin.

Schweizer, K., & Koch, W. (2002). A revision of Cattell's investment theory: Cognitive properties influencing learning. *Learning and Individual Differences, 13*, 57–82. https://doi.org/10.1016/S1041-6080(02)00062-6.

Sulkunen, S., & Malin, A. (2018). Literacy, age and recentness of education among Nordic adults. *Scandinavian Journal of Educational Research, 62*, 929–948. https://doi.org/10.1080/00313831.2017.1324898.

Tucker-Drob, E. M. (2011). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology, 47*, 331–343. https://doi.org/10.1037/a0021361.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122*, 231–249. https://doi.org/10.1037/0033-2909.122.3.231.

Vernon, P. (1965). Ability factors and environmental influences. *American Psychologist, 20*, 723–733. https://doi.org/10.1037/h0021472.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., … Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. *Psychological Science, 10*, 119–125. https://doi.org/10.1111/1467-9280.00118.

Zabal, A., Martin, S., & Rammstedt, B. (2017). *PIAAC-L data collection 2015: Technical report. GESIS papers 2017|29*. Cologne: GESIS.

Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and Aging, 17*, 690–695. https://doi.org/10.1037/0882-7974.17.4.690.