

RESEARCH ARTICLE

Measurement Invariance and Latent Mean Differences in the Reynolds Intellectual Assessment Scales (RIAS): Does the German Version of the RIAS Allow a Valid Assessment of Individuals with a Migration Background?

Jasmin T. Gygi*, Elodie Fux, Alexander Grob, Priska Hagmann-von Arx

Department of Psychology, University of Basel, Basel, Switzerland

* jasmin.gygi@unibas.ch



OPEN ACCESS

Citation: Gygi JT, Fux E, Grob A, Hagmann-von Arx P (2016) Measurement Invariance and Latent Mean Differences in the Reynolds Intellectual Assessment Scales (RIAS): Does the German Version of the RIAS Allow a Valid Assessment of Individuals with a Migration Background? PLoS ONE 11(11): e0166533. doi:10.1371/journal.pone.0166533

Editor: Pietro Cipresso, IRCCS Istituto Auxologico Italiano, ITALY

Received: June 27, 2016

Accepted: October 31, 2016

Published: November 15, 2016

Copyright: © 2016 Gygi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

This study examined measurement invariance and latent mean differences in the German version of the Reynolds Intellectual Assessment Scales (RIAS) for 316 individuals with a migration background (defined as speaking German as a second language) and 316 sex- and age-matched natives. The RIAS measures general intelligence (single-factor structure) and its two components, verbal and nonverbal intelligence (two-factor structure). Results of a multi-group confirmatory factor analysis showed scalar invariance for the two-factor and partial scalar invariance for the single-factor structure. We conclude that the two-factor structure of the RIAS is comparable across groups. Hence, verbal and nonverbal intelligence but not general intelligence should be considered when comparing RIAS test results of individuals with and without a migration background. Further, latent mean differences especially on the verbal, but also on the nonverbal intelligence index indicate language barriers for individuals with a migration background, as subtests corresponding to verbal intelligence require higher skills in German language. Moreover, cultural, environmental, and social factors that have to be taken into account when assessing individuals with a migration background are discussed.

Introduction

Migration has increased in recent decades in Europe [1]. For German-speaking countries, for example, the latest numbers reveal that 25% of the Swiss population [2] and 21% of the German population [3] have a migration background. The majority of immigrants originated from other European countries, including, for example, Italy and Turkey, while a smaller percentage originated from non-Western countries. The immigration of non-German-speaking individuals into Switzerland and Germany leads to more people seeking psychological assessment in a foreign-language environment in the respective countries.

Intelligence is one of the constructs most often studied in psychological practice [4]. Yet in individuals with a migration background, language barriers can lead to difficulties during assessment [5–7] such as misunderstanding of instructions or difficulties in articulating a verbal response. Thus, lower test scores in individuals with a migration background are particularly evident in tasks with high language requirements and their test performance increases with decreasing language dependence of the tasks [6–9]. Moreover, cultural and environmental factors may adversely affect test performance of individuals with a migration background [10]. For example, individuals with a migration background may be less familiar with the type of tasks used in performance assessment [11, 12]. In addition, many immigrant groups are confronted with negative achievement stereotypes, which may create increased performance pressure [13]. Furthermore, studies have shown that immigrant children's parents who have limited language skills in the dominant language are less involved in schooling, which in turn may negatively affect their children's performance [14, 15]. Consequently, in intelligence tests, individuals with a migration background may not be able to show their full potential, which can result in test scores that underestimate intelligence [5–7, 11, 12]. As high-stakes decisions are made on the basis of intelligence test scores, underestimated intelligence test scores may have negative consequences in, for example, education and employment. Regarding school-related decisions, underestimated intelligence test scores may lead to erroneous placement in special education services [16, 17]. In workplace situations, negative stereotypes may negatively influence performance ratings [18] and underestimated intelligence test scores may lead to not getting hired (e.g., [4]). Understanding how migration status influences test scores is therefore crucial to correctly interpreting the results.

To assess test score differences in individuals with and without a migration background, it is essential to examine measurement invariance prior to any interpretation of group differences [19]. Measurement invariance refers to the assumption of comparable relationships between items and their respective latent variables across groups [20]. When measurement invariance holds, researchers can validly compare statistical results, such as latent means, across these groups [21].

One currently used intelligence test is the Reynolds Intellectual Assessment Scales (RIAS) [22]. The RIAS is an individually administered test for individuals aged 3 to 90-plus years developed in the United States and recently adapted to different language groups including Danish [23], German [24], and Spanish [25]. The RIAS is designed to measure general intelligence and its two components, verbal and nonverbal intelligence. Additionally, the RIAS provides a measure of memory. Performance on the RIAS has been shown to be independent of motor coordination, visual-motor speed, and reading skills. Also, its administration, scoring, and interpretation have been described as user friendly [26–28]. However, we know of no study that has examined the RIAS intelligence factor structure across individuals with and without a migration background and whether these groups of individuals achieve comparable RIAS intelligence test scores.

The intelligence factor structure of the RIAS is based on Carroll's [29] three-stratum theory. On Stratum 1 the RIAS comprises two verbal subtests (Guess What, Verbal Reasoning) and two nonverbal subtests (Odd-Item Out, What's Missing?). On Stratum 2 the subtests are combined to create two factors, the Verbal Intelligence Index (VIX) and the Nonverbal Intelligence Index (NIX), which serve as indicators of crystallized and fluid intelligence, respectively. Stratum 3 comprises a Composite Intelligence Index (CIX), which is obtained by combining the VIX and NIX and which reflects general intelligence, *g*. A Composite Memory Index (CMX) on Stratum 2, which consists of two supplemental Stratum 1 subtests, is not integrated into the measure of general intelligence. On the basis of these theoretical assumptions, Reynolds and Kamphaus [22] suggested a two-factor intelligence structure with four subtests measuring two

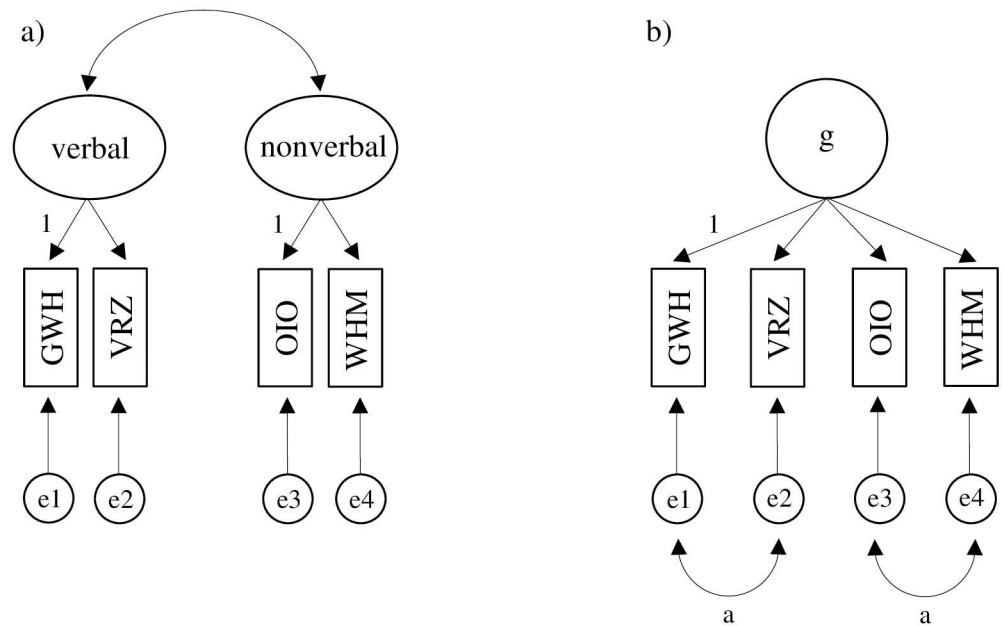


Fig 1. The Two-factor Structure (a) and Single-factor Structure (b) of the RIAS. *g* = general intelligence; GWH = Guess What; VRZ = Verbal Reasoning; OIO = Odd-Item Out; WHM = What’s Missing; e1-e4 = residuals; a = covariances of the residuals were set to be equal across groups.

doi:10.1371/journal.pone.0166533.g001

factors, verbal and nonverbal, (Fig 1a). Another structure assessed by Reynolds and Kamphaus [22] was a single-factor structure with four subtests measuring general intelligence (Fig 1b).

The RIAS intelligence factor structure proposed by Reynolds and Kamphaus [22] (Fig 1a) has been supported in the English, Danish, German, and Spanish RIAS versions by using confirmatory factor analysis (CFA) and standardization samples [22–25] as well as referred students in the English version [30]. Further, studies using CFA procedures also found support for the single-factor structure in the standardization samples of the four RIAS versions [31] as well as in typically developing and referred samples in the English version [32]. Additionally, studies using exploratory factor analysis have found evidence for a single-factor structure (Fig 1b) for the English RIAS standardization sample [33] and for referred samples [34–36]. While the theoretically proposed RIAS intelligence factor structure was supported in standardization and referred samples, it is unknown whether there is measurement invariance across groups with and without a migration background.

The main goal of the present study was to assess measurement invariance and mean differences in the German version of the RIAS in individuals with and without a migration background. The German version of the RIAS was standardized in Switzerland and Germany. Both countries have a high percentage of individuals with a migration background of over 20% of the population [2, 3]. Therefore, it is essential to ensure measurement invariance of diagnostic instruments such as the RIAS, so that examiners can rely on the validity of test results for individuals with a migration background. In the current study we addressed the following research questions: (a) Are the intelligence factor structures (i.e., single-factor structure and two-factor structure) of the RIAS invariant across individuals with and without a migration background, and (b) do individuals with and without a migration background differ in latent means and variances in the RIAS intelligence factors (the VIX, NIX, and CIX)?

First, we conducted single-group CFAs to determine model fit for the RIAS intelligence factor structures depicted in Fig 1a and 1b for both groups separately. Second, we ran a multi-

group CFA for both RIAS intelligence factor structures to assess measurement invariance across groups. We hypothesized that measurement invariance would hold across groups. When invariance held, we examined invariance of the latent variances and latent means across groups. We assumed that individuals with a migration background would have lower means in the latent variables VIX, NIX, and CIX than individuals without a migration background. We further assumed that lower latent means would in particular be evident in the RIAS VIX, drawing on the assumption that language barriers might have led to lower test performance as the verbal subtests require higher skills in German language than the nonverbal subtests [6–9].

Method

Participants

The study included 632 individuals with and without a migration background (each $n = 316$) from the standardization sample of the German version of the RIAS (standardization sample $N = 2,145$) [24]. This study was carried out in accordance with the recommendations of the Declaration of Helsinki, and it was approved by the Ethics Committee of Basel. Written informed consent was obtained from all subjects. For the children, parents gave written informed consent and assent was obtained from each child prior to the start of the study. Individuals with and without a migration background were matched in sex and age. Sex was equally distributed, with 48% female and 52% male, according to chi-square tests ($p > .05$). Subjects were between the ages of 3 and 99 years ($M = 15.79$, $SD = 16.81$). The frequencies of individuals in different age ranges are provided in [S1 Table](#). Migration background was defined as speaking German as a first or a second language: Individuals without a migration background all spoke German as their first language, whereas individuals with a migration background all spoke German as a second language. In our study, 88% of individuals with a migration background spoke a language from a Western, European country as their first language, while 12% of the immigrants spoke a language from a non-Western country as their first language. First languages most often named by individuals with a migration background were Turkish (21%), Italian (8%), Serbian (6%), and Spanish (5%). [S2 Table](#) gives a detailed overview of all foreign first languages. The distribution of Western and non-Western immigrants represents the immigrant population in Switzerland and Germany well [2, 3], and thus all individuals with a migration background were included in the following analyses. However, there are studies that revealed group differences in intelligence across individuals from Western and immigrants from non-Western countries [37]. In the present study, the sample of immigrants from non-Western countries was not large enough to analyze such group differences. However, we also performed all analyses with immigrants from Western European countries only. As the results remained stable, we present results only for the whole sample including immigrants from both Western and non-Western countries, as we believe this distribution is more representative of the immigrant populations in Switzerland and Germany.

Measures

The RIAS [22] is an individually administered intelligence test. Its German version [24] was standardized for individuals aged 3 to 99 years in Germany and Switzerland ($N = 2,145$) in 2011 and 2012. The RIAS is composed of four intelligence subtests, which together constitute the CIX. The CIX can be further divided into two indices, represented by two subtests each: The verbal, VIX, comprises the subtests Guess What (GWH), in which individuals are asked to identify an object or a concept through the use of verbally presented clues, and Verbal Reasoning (VRZ), in which individuals are asked to complete verbal analogies. Therefore, the verbal subtests require high language skills. The nonverbal, NIX, comprises the subtests Odd-Item

Out (OIO), in which individuals are asked to identify a picture that does not go with others, and What's Missing? (WHM), in which individuals are required to identify the missing element in a presented picture. The CIX is based on the sum of the *T* scores ($M = 50$; $SD = 10$) of the four intelligence subtests. Reliability for the German RIAS is high for both groups with Cronbach's $\alpha > .90$ for the subtests and $\alpha > .95$ for intelligence indexes.

Data analysis

To evaluate measurement invariance and invariance of latent variances and means across individuals with and without a migration background we used SPSS AMOS version 22 [38]. Correlation matrices with means and standard deviations of the subtest *T* scores were used as the input data file (see Table 1).

Following the suggestion of Meade, Johnson, and Braddy [39], we first conducted single-group CFAs to test the two-factor and single-factor structures, depicted in Fig 1, for each group. Second, we conducted a multi-group CFA for both factor structures across the two groups. Different levels of measurement invariance were tested, as described by Meredith [20], Milfont and Fischer [40], Steenkamp and Baumgartner [41], and Widaman and Reise [19]. First, we assessed configural invariance (Model 1). The intercepts of GWH and OIO were fixed to zero across both groups for the two-factor structure, and the intercept of GWH was

Table 1. Correlations Between Reynolds Intellectual Assessment Scales Subtests and Indexes for Individuals With and Without a Migration Background With Means and Standard Deviations.

Subtest or index	GWH ^a	VRZ ^a	OIO ^a	WHM ^a	VIX ^b	NIX ^b	CIX ^b
Without migration background							
GWH	1.00						
VRZ	.58	1.00					
OIO	.30	.35	1.00				
WHM	.32	.28	.38	1.00			
VIX	.88	.90	.36	.33	1.00		
NIX	.37	.38	.83	.83	.42	1.00	
CIX	.75	.77	.70	.68	.86	.83	1.00
Mean	53.11	53.39	52.21	51.60	105.85	103.50	105.34
SD	8.48	9.61	8.85	8.75	13.52	12.46	12.43
With migration background							
GWH	1.00						
VRZ	.65	1.00					
OIO	.37	.43	1.00				
WHM	.31	.32	.57	1.00			
VIX	.93	.88	.43	.34	1.00		
NIX	.38	.42	.89	.88	.44	1.00	
CIX	.77	.77	.78	.72	.85	.85	1.00
Mean	44.29	45.58	47.99	47.61	91.86	96.54	93.44
SD	11.28	8.96	10.66	10.28	15.39	15.68	14.92
Cohen's <i>d</i> ^c	1.04	0.81	0.48	0.46	1.03	0.56	0.96

Note. $N_{\text{without migration background}} = 316$, $N_{\text{with migration background}} = 316$. *SD* = standard deviation; GWH = Guess What; VRZ = Verbal Reasoning; OIO = Odd-Item Out; WHM = What's Missing; VIX = Verbal Intelligence Index; NIX = Nonverbal Intelligence Index; CIX = Composite Intelligence Index.

^a*T* score normative mean = 50, *SD* = 10.

^bIntelligence indexes normative mean = 100, *SD* = 15.

^cCohen's *d* is calculated based on the manifest data.

doi:10.1371/journal.pone.0166533.t001

fixed to zero across both groups for the single-factor structure. Second, we tested metric invariance (known as weak invariance; Model 2) by additionally constraining factor loadings across groups. Third, we examined scalar invariance (known as strong invariance; Model 3), by additionally constraining intercepts across groups. Fourth, we assessed residual invariance (known as strict invariance; Model 4) by additionally constraining error variances across groups. Finally, when scalar invariance held, invariance of the latent factors' variances and means could be analyzed [19] by constraining factor variances (Model 5) and factor means (Model 6) using the scalar invariance model as baseline.

For the single-factor structure, following Beaujean and McGlaughlin [34], we let the residuals of GWH and VRZ as well as the residuals of OIO and WHM to covary and constrained these two covariances to be equal [36]. We stress that this structure leads to a model empirically equivalent [42] to the two-factor structure as well as to a bifactor structure with a single general factor and two specific factors each measured by two indicators with factor loadings fixed to 1 [43, 44].

Following Hu and Bentler [45], we used a Comparative Fit Index (CFI) of $\geq .95$, a McDonald's Noncentrality Index (Mc) of $\geq .90$, and a root mean square error of approximation (RMSEA) of $\leq .06$ to evaluate good model fit. Because chi-square test statistics tend to be sensitive to large sample sizes [39], we relied on the test of small difference in fit with $p < .05$ [46] and changes in goodness-of-fit indices with $\Delta CFI \leq .002$ and $\Delta Mc \leq .005$ [39] to compare the nested models. We accepted model invariance if two of the three difference in fit statistics were within cutoff points.

Additionally, we computed Cohen's d [47] to investigate the effect sizes of the manifest and latent mean differences. A value of $d = 0.2$ is considered a small effect, $d = 0.5$ a medium effect, and $d = 0.8$ a large effect.

Results

Single-group CFA

The two-factor (Fig 1a) and single-factor (Fig 1b) RIAS intelligence structures were evaluated for both groups using single-group CFA. As can be seen in Table 2, both structures yielded a good model fit for each group (CFI $\geq .993$, Mc $\geq .998$, RMSEA $\leq .070$), indicating the models to be empirically equivalent [42]. Therefore, both structures were further analyzed to assess invariance of the latent variances and means of verbal and nonverbal intelligence (i.e., for this we analyzed the two-factor structure), as well as general intelligence (i.e., for this we analyzed

Table 2. Fit Indices for Single-group Confirmatory Factor Analyses Evaluating Two Reynolds Intellectual Assessment Scales Intelligence Factor Structures in Individuals With and Without a Migration Background.

Structure	df ⁰	χ^2	CFI	Mc	RMSEA	90% CI
Two factors ^a						
Without migration background	1	2.556	.993	.998	.070	[.000, .183]
With migration background	1	0.438	1.000	1.000	< .001	[.000, .129]
Single factor ^a						
Without migration background	1	2.556	.993	.998	.070	[.000, .183]
With migration background	1	0.438	1.000	1.000	< .001	[.000, .129]

Note. $N_{\text{without migration background}} = 316$, $N_{\text{with migration background}} = 316$. CFI = Comparative Fit Index; Mc = McDonald's Noncentrality Index; RMSEA = root mean square error of approximation; CI = confidence interval.

^aIncludes four subtests.

doi:10.1371/journal.pone.0166533.t002

the single-factor structure). The factor loadings, covariances, and correlations for both RIAS structures are provided in S3 Table.

Multi-group CFA

Measurement invariance across both groups was analyzed using multi-group CFA. For the two-factor structure (Fig 1a) results are shown in Table 3. Configural, metric, and scalar invariance did hold across both groups with regard to ΔCFI, ΔMc, and the *p* value of the test of small difference in fit. Residual invariance yielded a poor fit regarding ΔCFI and ΔMc, though the *p* value of the test of small difference in fit was < .05. Hence, we conclude that the variances of the residuals were noninvariant across groups.

Further, invariance of the latent variances and means for verbal and nonverbal intelligence (two-factor structure) were tested using the model implying scalar invariance as baseline (see Table 3). On the level of scalar invariance, results regarding verbal intelligence indicate that latent variances were invariant between groups with $s^2 = 117.20$ for individuals with a migration background and $s^2 = 72.00$ for individuals without a migration background. Latent means differed between groups and were lower for individuals with a migration background ($M = 44.44$) compared to individuals without a migration background ($M = 53.05$) with an effect size of $d = 1.01$, indicating a large effect. Regarding nonverbal intelligence, latent variances and latent means differed between groups. Latent variances were higher and latent means lower in individuals with a migration background ($s^2 = 107.40$, $M = 47.91$) compared to individuals without a migration background ($s^2 = 72.81$, $M = 52.30$). According to Cohen [47], the effect size of the group difference for nonverbal intelligence was medium with $d = 0.52$.

Results of the single-factor structure (Fig 1b) are presented in Table 4. Though configural and metric invariance held, the models implying scalar and residual invariance showed values of ΔCFI and ΔMc above the suggested cutoff values. Although the *p* value of the test of small

Table 3. Fit Indices for Multi-group Confirmatory Factor Analysis Evaluating Measurement Invariance of the Two-factor Structure Including the Four Intelligence Subtests Across Individuals With and Without a Migration Background.

Model	df	χ^2	CFI	Mc	RMSEA	90% CI	Δdf	$\Delta\chi^2$	ΔCFI	ΔMc	Small diff <i>p</i>
1 Configural invariance	2	2.994	.998	.999	.028	[.000, .089]	–	–	–	–	–
2 Metric invariance	4	6.010	.997	.998	.028	[.000, .071]	–	–	–	–	–
2 versus 1	–	–	–	–	–	–	2	3.016	.001	< .001	.221
3 Scalar invariance	6	7.309	.998	.999	.019	[.000, .057]	–	–	–	–	–
3 versus 2	–	–	–	–	–	–	2	1.299	< .001	< .001	.522
4 Residual invariance	10	37.877	.954	.978	.067	[.045, .090]	–	–	–	–	–
4 versus 3	–	–	–	–	–	–	4	30.567	.044	.021	.196
5a Factor variances verbal ^a	7	12.273	.991	.996	.035	[.000, .066]	–	–	–	–	–
5a versus 3	–	–	–	–	–	–	1	4.963	.007	.003	.296
5b Factor variances nonverbal ^a	7	24.161	.971	.986	.062	[.036, .090]	–	–	–	–	–
5b versus 3	–	–	–	–	–	–	1	16.852	.027	.013	.217
6a Factor means verbal ^a	7	139.309	.780	.900	.173	[.149, .199]	–	–	–	–	–
6a versus 3	–	–	–	–	–	–	1	132.000	.218	.099	.031
6b Factor means nonverbal ^a	7	43.862	.939	.971	.091	[.067, .118]	–	–	–	–	–
6b versus 3	–	–	–	–	–	–	1	36.553	.059	.028	.145

Note. $N_{Total} = 632$, $N_{Without migration background} = 316$, $N_{With migration background} = 316$. CFI = Comparative Fit Index; Mc = McDonald’s Noncentrality Index; RMSEA = root mean square error of approximation; CI = confidence interval.

^aFactor variance invariance and factor mean invariance were calculated at the level of scalar invariance.

doi:10.1371/journal.pone.0166533.t003

Table 4. Fit Indices for Multi-group Confirmatory Factor Analysis Evaluating Measurement Invariance of the Single-factor Structure Including the Four Intelligence Subtests Across Individuals With and Without a Migration Background.

Model	df	χ^2	CFI	Mc	RMSEA	90% CI	Δdf	$\Delta \chi^2$	ΔCFI	ΔMc	Small diff <i>p</i>
1 Configural invariance	3	11.506	.986	.993	.067	[.029, .110]	–	–	–	–	–
2 Metric invariance	6	16.481	.983	.992	.053	[.023, .084]	–	–	–	–	–
2 versus 1	–	–	–	–	–	–	3	4.975	.003	.001	.174
3 Scalar invariance	9	29.839	.965	.984	.061	[.037, .085]	–	–	–	–	–
3 versus 2	–	–	–	–	–	–	3	13.358	.018	.008	.271
4 Residual invariance	13	62.525	.918	.962	.078	[.059, .098]	–	–	–	–	–
4 versus 2	–	–	–	–	–	–	4	32.686	.047	.022	.181
5 Factor variances ^a	10	38.824	.952	.977	.068	[.046, .091]	–	–	–	–	–
5 versus 2	–	–	–	–	–	–	1	8.985	.013	.007	.270
6 Factor means ^a	10	150.616	.766	.895	.149	[.129, .171]	–	–	–	–	–
6 versus 2	–	–	–	–	–	–	1	120.777	.199	.089	.034

Note. $N_{Total} = 632$, $N_{Without\ migration\ background} = 316$, $N_{With\ migration\ background} = 316$. CFI = Comparative Fit Index; Mc = McDonald’s Noncentrality Index; RMSEA = root mean square error of approximation; CI = confidence interval.

^a Factor variance invariance and factor mean invariance were calculated at the level of measurement invariance, as scalar invariance did not hold.

doi:10.1371/journal.pone.0166533.t004

difference in fit is $< .05$, we conclude that there are differences in subtests’ intercepts and residual variances across groups. As scalar invariance did not hold, we examined partial scalar invariance to identify specific subtests that caused noninvariance on the scalar level as proposed by Steenkamp and Baumgartner [41]. We found that the subtest OIO was the major contributor to the worsened model fit (S4 Table). Variances and means of general intelligence were not analyzed due to the lack of full scalar invariance [20].

Discussion

The main goal of this study was to assess measurement invariance and differences in latent means of the German version of the RIAS in individuals with and without a migration background. We analyzed a two-factor structure with verbal and nonverbal intelligence as latent factors as well as a single-factor structure with general intelligence as the latent factor.

Both the two-factor and the single-factor structure were supported through single-CFAs for each group separately. This is in line with the RIAS technical manual [22, 24] and with previous research supporting the two-factor structure [30], the single-factor structure [23–26], or both RIAS intelligence factor structures [31, 32].

Then, we conducted multi-group CFAs to analyze measurement invariance for both structures across individuals with and without a migration background. For the two-factor structure, results showed scalar invariance across groups, indicating that differences in the subtest means were due to differences in the means of the underlying constructs [48]. Residual invariance could not be supported, possibly because of differing reliabilities of the scales across groups [49].

Further, we assessed invariance of latent variances and means of verbal and nonverbal intelligence. Results indicate that variances of nonverbal intelligence as well as means of verbal and nonverbal intelligence were noninvariant across groups. Individuals with a migration background showed similar variances in verbal intelligence and higher variances in nonverbal intelligence compared to individuals without a migration background. These results suggest that individuals with a migration background showed about the same range in verbal intelligence test scores but a wider range in nonverbal intelligence test scores than individuals

without a migration background. Beaujean and McGlaughlin [34] found a similar pattern when examining referred black and white students, but black students showed a higher variance in general intelligence in the RIAS test scores. Further, as hypothesized, individuals with a migration background achieved lower latent means in verbal as well as nonverbal intelligence when tested with the German version of the RIAS. This is in line with other studies that also found lower means in intelligence test scores for individuals with a migration background [5, 7, 9]. For the group differences, the effect size in verbal intelligence was high, whereas the effect size in nonverbal intelligence was medium, according to Cohen [47]. This is in accordance with our hypothesis that differences in latent means would be especially evident in the RIAS VIX. A possible explanation for this result may be that language barriers in individuals with a migration background might have led to lower test performance especially in the RIAS subtests with high language requirements [6, 8, 9]. However, individuals with a migration background also showed lower test performance in nonverbal intelligence. Existing literature suggests that not only language barriers but also cultural and environmental differences such as familiarity with tasks used in intelligence tests [11, 12], stereotype threat [13, 18], and parental language skills [14] can contribute to lower test performance in individuals with a migration background [10], which also might have led to the lower test scores of individuals with a migration background in the verbal and nonverbal intelligence indexes in the present study. In sum, the results of the present study indicate that using the German version of the RIAS may lead to underestimated intelligence test scores in individuals with a migration background. Such an approach in which examiners focus exclusively on the test results is called a “static” intelligence assessment ([12], p. 445). In contrast, “dynamic” testing methods ([12], p. 445) may constitute an assessment approach that is able to account for differences in individuals from different cultural contexts. In line with the zone of proximal development concept [50], in dynamic testing, the tasks are first introduced to familiarize the individuals with the nature of these tasks. Afterward the improvement in these tasks is assessed to evaluate the learning potential of an examinee, which may be a better estimate of abilities compared to results gained from a static assessment [51]. Thus, there is evidence that dynamic testing leads to diminishing test score differences across individuals with and without a migration background [11, 12]. These results indicate that examiners in psychological practice should consider using a dynamic testing approach to reduce effects of familiarity with intelligence test tasks. It should be a goal for future studies to determine whether the positive effects of such a dynamic approach are also evident when using the German version of the RIAS in individuals with and without a migration background.

For the single-factor structure, only metric invariance and not scalar invariance held. This indicates that the relationship between the observed scores and the latent construct was different across groups [20]. Additionally, we assessed partial scalar invariance and found that the major contributor to differences across groups was the subtest OIO. This is in line with other studies that also found OIO to be the major contributor to a lack of full scalar invariance [30] or a decreased model fit of scalar invariance [34]. Thus, group differences in the latent means can be explained either through the mean differences in the subtest OIO or through true latent mean differences [48].

Because full scalar invariance did not hold for the single-factor structure, we did not analyze invariance of latent variances and means of general intelligence [21, 41]. This leads to the conclusion that test scores of individuals with a migration background should not be compared with test scores of individuals without a migration background on the level of general intelligence when tested with the RIAS, but rather on the level of verbal and nonverbal intelligence only.

There are strengths and limitations of this study. We see it as a strength that we analyzed competing factor structures (two-factor vs. single-factor) of the RIAS, and thus our results contribute to the discussion on the best-fitting RIAS intelligence factor structure for specific samples. Further, we analyzed large age- and sex-matched samples of individuals with and without a migration background. However, we could not control for social factors such as socioeconomic status (SES); previous studies have shown that SES is positively correlated with intelligence [52] and negatively related to migration background such that individuals with a migration background have a lower SES than individuals without a migration background [53, 54]. Therefore, it is possible that the means in individuals with a migration background would have been slightly higher if we additionally could have controlled for SES. Further, we determined migration background solely on the basis of individuals being foreign-language or native. Future studies could additionally gather the status of migration (country of birth, number of the family's generations living in the country of immigration, etc.). Finally, our sample of individuals with a migration background included subjects with a diversity of first languages. Although this reflects the language diversity of Switzerland and Germany [2, 3], it can be assumed that these individuals do not represent a homogeneous group. Future studies might therefore focus on specific groups such as individuals from countries with collectivist or individualist cultural values [55] or from different language families [56]. Future studies might also analyze RIAS test scores in individuals with and without a migration background in other countries with a high percentage of migration.

Conclusion

In conclusion, the RIAS two-factor intelligence structure including four subtests measuring verbal and nonverbal intelligence showed measurement invariance across samples, whereas we could only demonstrate partial scalar invariance for the single-factor structure with four subtests measuring general intelligence. These results suggest that examiners should focus on the VIX and NIX and not on the CIX when comparing RIAS test results of individuals with and without a migration background. Finally, examiners should consider the migration background as well as cultural, environmental, and social factors of an examinee when interpreting RIAS test results, as individuals with a migration background may achieve lower test scores in particular on the VIX but also on the NIX than individuals without a migration background, which may be due to language, cultural, environmental, or social barriers.

Supporting Information

S1 Table. Frequencies for Age Ranges for the Total Sample.

(DOCX)

S2 Table. Overview of All First Languages Named by Subjects With a Migration Background.

(DOCX)

S3 Table. Unstandardized and Standardized Factor Loadings, Covariances, and Correlations for the Two-Factor and Single-Factor Structures of the Confirmatory Factor Analyses for Individuals With and Without a Migration Background.

(DOCX)

S4 Table. Fit Indices for Multi-group Confirmatory Factor Analysis Evaluating Partial Measurement Invariance of the Single-factor Structure Including the Four Intelligence Subtests Across Individuals With and Without a Migration Background.

(DOCX)

Author Contributions

Conceptualization: JTG EF AG PH.

Data curation: JTG EF PH.

Formal analysis: JTG EF PH.

Funding acquisition: AG PH.

Investigation: JTG EF AG PH.

Methodology: JTG EF PH AG.

Project administration: AG PH.

Resources: AG PH.

Supervision: PH.

Visualization: JTG.

Writing – original draft: JTG PH.

Writing – review & editing: JTG EF AG PH.

References

1. Eurostat. The EU in the world—population; 2014. Available: http://ec.europa.eu/eurostat/statistics-explained/index.php/The_EU_in_the_world_-_population#Migration_and_asylum
2. Bundesamt für Statistik. Migration und Integration; 2015. Available: <http://www.bfs.admin.ch/bfs/portal/de/index/themen/01/07/blank/key/01/01.htmlx>
3. Statistisches Bundesamt. Bevölkerungsstand; 2015. Available: <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/Bevoelkerungsstand/Bevoelkerungsstand.html;jsessionid=D9339CCA1586CB4B47CA4D661CFD1EE8.cae1>
4. Goldstein S, Princiotta D, Naglieri JA, editors. Handbook of intelligence. Evolutionary theory, historical perspective, and current concepts. New York, NY: Springer; 2015. doi: [10.1007/978-1-4939-1562-0](https://doi.org/10.1007/978-1-4939-1562-0)
5. Calero MD, Fernández-Parra A, López-Rubio S, Carles R, Mata S, Vives C, et al. Variables involved in personal, social and school adjustment in a sample of preschool-aged children from different cultural backgrounds. *Eur J Psychol Educ*. 2013; 28:133–155. doi: [10.1007/s10212-012-0107-8](https://doi.org/10.1007/s10212-012-0107-8)
6. Daseking M, Lipsius M, Petermann D, Waldmann H-C. Application of the German WISC-IV in children with diverse cultural and linguistical backgrounds. *Kindheit und Entwicklung*. 2008; 17:76–89.
7. Hagmann-von Arx P, Petermann F, Grob A. Konvergente und diskriminante Validität der WISC-IV und der Intelligence and Development Scales (IDS) bei Kindern mit Migrationshintergrund [Convergent and discriminant validity of the WISC-IV and the Intelligence and Development Scales (IDS) in children with a migration background]. *Diagnostica*. 2013; 59:170–182. doi: [10.1026/0012-1924/a000091](https://doi.org/10.1026/0012-1924/a000091)
8. Weiss LG., Harris JG, Prifitera A, Courville T, Rolfhus E, Saklofske DH, et al. WISC-IV interpretation in societal context. In: Weiss LG, Saklofske DH, Prifitera A, Holdnack JA, editors. *WISC-IV advanced clinical interpretation*. San Diego, CA: Elsevier; 2006. pp. 1–57.
9. Harris JG, Muñoz MR, Llorente AM. Considerations in using the WISC-IV with Hispanic children. In: Prifitera A, Saklofske DH, Weiss LG, editors. *WISC-IV Clinical assessment and intervention*. San Diego, CA: Elsevier; 2008. pp. 409–442.
10. Mackintosh NJ. *IQ and human intelligence*. 2nd ed. New York: Oxford University Press; 2011.
11. Calero MD, Mata S, Carles R, Vives C, López-Rubio S, Fernández-Parra A, et al. Learning potential assessment and adaptation to the educational context: The usefulness of the ACFS for assessing immigrant preschool children. *Psychol Sch*. 2013; 50:705–721. doi: [10.1002/pits.21701](https://doi.org/10.1002/pits.21701)
12. Resing WC, Tunteler E, de Jong FM, Bosma T. Dynamic testing in indigenous and ethnic minority children. *Learn Individ Differ*. 2009; 19:445–450. doi: [10.1016/j.lindif.2009.03.006](https://doi.org/10.1016/j.lindif.2009.03.006)
13. Appel M, Weber S, Kronberger N. The influence of stereotype threat on immigrants: review and meta-analysis. *Frontiers Psychol*. 2015; 6:900. doi: [10.3389/fpsyg.2015.00900](https://doi.org/10.3389/fpsyg.2015.00900) PMID: [26217256](https://pubmed.ncbi.nlm.nih.gov/26217256/)

14. Lahaie C. School readiness of children of immigrants: Does parental involvement play a role? *Soc Sci Q.* 2008; 89:684–705.
15. Turney K, Kao G. Barriers to school involvement: Are immigrant parents disadvantaged? *J Educ Res.* 2009; 102:257–271. doi: [10.3200/JOER.102.4.257-271](https://doi.org/10.3200/JOER.102.4.257-271)
16. Hessels MG. Low I.Q. but high learning potential: Why Zeineb and Moussa do not belong in special education. *Educ Child Psychol.* 1997; 14:121–136.
17. Klingner JK, Blanchett WJ, Harry B. Race, culture, and developmental disabilities. In: Odom SL, Horner RH, Snell ME, editors. *Handbook of developmental disabilities.* New York: Guilford Press; 2007. pp. 55–75.
18. Baltés BB, Rudolph CW. Examining the effect of negative Turkish stereotypes on evaluative workplace outcomes in Germany. *J Manag Psychol.* 2010; 25:148–158. doi: [10.1108/02683941011019357](https://doi.org/10.1108/02683941011019357)
19. Widaman KF, Reise SP. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In Bryant K, Windle M, West S, editors. *The science of prevention: Methodological advances from alcohol and substance abuse research.* Washington, DC: American Psychological Association; 1997. pp. 281–324. doi: [10.1037/10222-009](https://doi.org/10.1037/10222-009)
20. Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika.* 1993; 48:525–543. doi: [10.1007/BF02294825](https://doi.org/10.1007/BF02294825)
21. Chen FF, Sousa KH, West SG. Testing measurement invariance of second-order factor models. *Struct Equ Modeling.* 2005; 12:471–492.
22. Reynolds CR, Kamphaus RW. *Reynolds Intellectual Assessment Scales: Professional manual.* Lutz, FL: PAR; 2003.
23. Hartmann P, Andresen J. *Dokumentation af den danske standardisering af RIAS.* Virum, Denmark: Hogrefe Psykologisk; 2011.
24. Hagmann-von Arx P, Grob A. *Reynolds Intellectual Assessment Scales (RIAS).* Bern, Switzerland: Hans Huber; 2014.
25. Santamaría P, Fernández Pinto I. *RIAS Escalas de Inteligencia de Reynolds y RIST Test de Inteligencia Breve de Reynolds Manual. Adaptación española.* Madrid, Spain: TEA Ediciones; 2009.
26. Andrews JJW. RIAS: Reynolds Intellectual Assessment Scales. *J Psychoeduc Assess.* 2007; 25:402–408.
27. Dombrowski SC, Mrazik M. Test review: Reynolds Intellectual Assessment Scales. *Can J Sch Psychol.* 2008; 23:223–230.
28. Elliot RW. Test review: Reynolds Intellectual Assessment Scales. *Arch Clin Neuropsychol.* 2004; 19:325–328. PMID: [15046017](https://pubmed.ncbi.nlm.nih.gov/15046017/)
29. Carroll JB. *Human cognitive abilities: A survey of factor-analytic studies.* New York, NY: Cambridge University Press; 1993.
30. Beaujean AA, McGlaughlin SM, Margulies AS. Factorial validity of the Reynolds Intellectual Assessment Scales for referred students. *Psychol Sch.* 2009; 46:932–950. doi: [10.1002/pits.20435](https://doi.org/10.1002/pits.20435)
31. Gygi JT, Ledermann T, Grob A, Hagmann-von Arx P. The Reynolds Intellectual Assessment Scales (RIAS): Measurement invariance across four language groups. Manuscript submitted for publication; 2016.
32. Irwin JK, Joschko M, Kerns KA. Confirmatory factor analysis of the Reynolds Intellectual Assessment Scales (RIAS) in Canadian children. *Clin Neuropsychol.* 2014; 28:1258–1277. doi: [10.1080/13854046.2014.975843](https://doi.org/10.1080/13854046.2014.975843) PMID: [25359205](https://pubmed.ncbi.nlm.nih.gov/25359205/)
33. Dombrowski SC, Watkins MW, Brogan MJ. An exploratory investigation of the factor structure of the Reynolds Intellectual Assessment Scales (RIAS). *J Psychoeduc Assess.* 2009; 27:494–507. doi: [10.1177/0734282909333179](https://doi.org/10.1177/0734282909333179)
34. Beaujean AA, McGlaughlin SM. Invariance in the Reynolds Intellectual Assessment Scales for black and white referred students. *Psychol Assess.* 2014; 26:1394–1399. doi: [10.1037/pas0000029](https://doi.org/10.1037/pas0000029) PMID: [25285717](https://pubmed.ncbi.nlm.nih.gov/25285717/)
35. Nelson JM, Canivez GL. Examination of the structural, convergent, and incremental validity of the Reynolds Intellectual Assessment Scales (RIAS) with a clinical sample. *Psychol Assess.* 2012; 24:129–140. doi: [10.1037/a0024878](https://doi.org/10.1037/a0024878) PMID: [21823799](https://pubmed.ncbi.nlm.nih.gov/21823799/)
36. Nelson JM, Canivez GL, Lindstrom W, Hatt CV. Higher-order exploratory factor analysis of the Reynolds Intellectual Assessment Scales with a referred sample. *J Sch Psychol.* 2007; 45:439–456. doi: [10.1016/j.jsp.2007.03.003](https://doi.org/10.1016/j.jsp.2007.03.003)
37. te Nijenhuis J, Willigers D, Dragt J, van der Flier H. The effects of language bias and cultural bias estimated using the method of correlated vectors on a large database of IQ comparisons between native Dutch and ethnic minority immigrants from non-Western countries. 2016; 54:117–135.

38. Arbuckle JL. IBM SPSS Amos 22 user's guide. Armonk, NY: IBM; 2013.
39. Meade AW, Johnson EC, Braddy PW. Power and sensitivity of alternative fit indices in tests of measurement invariance. *J Appl Psychol*. 2008; 93:568–592. doi: [10.1037/0021-9010.93.3.568](https://doi.org/10.1037/0021-9010.93.3.568) PMID: [18457487](https://pubmed.ncbi.nlm.nih.gov/18457487/)
40. Milfont TL, Fischer R. Testing measurement invariance across groups: Applications in cross-cultural research. *Int J Psychol Res*. 2010; 3:111–130.
41. Steenkamp JBE, Baumgartner H. Assessing measurement invariance in cross-national consumer research. *J Consum Res*. 1998; 25:78–107. doi: [10.1086/209528](https://doi.org/10.1086/209528)
42. Lee S, Hershberger S. A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behav Res*. 1990; 25:313–334. doi: [10.1207/s15327906mbr2503_4](https://doi.org/10.1207/s15327906mbr2503_4) PMID: [26761406](https://pubmed.ncbi.nlm.nih.gov/26761406/)
43. Canivez GL, Watkins MW. Review of the Wechsler Intelligence Scale for Children—Fifth Edition: Critique, commentary, and independent analyses. In: Kaufman AS, Raiford SE, Coalson DL, editors. *Intelligent testing with the WISC–V*. Hoboken, NJ: Wiley; 2016. pp. 683–702.
44. Reise SP. The rediscovery of bifactor measurement models. *Multivariate Behav Res*. 2012; 47:667–696. doi: [10.1080/00273171.2012.715555](https://doi.org/10.1080/00273171.2012.715555) PMID: [24049214](https://pubmed.ncbi.nlm.nih.gov/24049214/)
45. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999; 6:1–55. doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
46. MacCallum RC, Browne MW, Cai L. Testing differences between nested covariance structure models: Power analysis and null hypotheses. *Psychol Methods*. 2006; 11:19–35. doi: [10.1037/1082-989X.11.1.19](https://doi.org/10.1037/1082-989X.11.1.19) PMID: [16594765](https://pubmed.ncbi.nlm.nih.gov/16594765/)
47. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum; 1988.
48. Meredith W, Teresi JA. An essay on measurement and factorial invariance. *Medical Care*. 2006; 44:69–77. doi: [10.1097/01.mlr.0000245438.73837.89](https://doi.org/10.1097/01.mlr.0000245438.73837.89) PMID: [17060838](https://pubmed.ncbi.nlm.nih.gov/17060838/)
49. DeShon RP. Measures are not invariant across groups without error variance homogeneity. *Psychol Sci*. 2004; 46:137–149.
50. Vygotsky LS. Interaction between learning and development (M. Lopez-Morillas, Trans.). In: Cole M, John-Steiner V, Scribner S, Souberman E, editors. *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press; 1978. pp. 79–91.
51. Sternberg RJ, Grigorenko EL. *Dynamic testing: The nature and measurement of learning potential*. New York: Cambridge University Press; 2002.
52. Bradley RH, Corwyn RF. Socioeconomic status and child development. *Annu Rev Psychol*. 2002; 53:371–399. doi: [10.1146/annurev.psych.53.100901.135233](https://doi.org/10.1146/annurev.psych.53.100901.135233) PMID: [11752490](https://pubmed.ncbi.nlm.nih.gov/11752490/)
53. Bundesamt für Statistik. Verteilung des verfügbaren Äquivalenzeinkommens und das Quintilverhältnis S80/S20, nach verschiedenen soziodemografischen Merkmalen; 2014. Available: <http://www.bfs.admin.ch/bfs/portal/de/index/themen/20/03/blank/key/02/06.html>.
54. DeNavas-Walt C, Proctor BD. Income and poverty in the United States: 2014. U.S. Census Bureau Current Population Reports, P60-252; 2015. Available: <http://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-252.pdf>
55. Hofstede G. *Culture's consequences comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage; 2001.
56. Ruhlen M. *A guide to the world's languages: Volume I, Classification*. Stanford, CA: Stanford University Press; 1991.