Validity study of intelligence subtests for a Brazilian giftedness assessment battery

Validade de subtestes de inteligência para a bateria de avaliação de altas habilidades/superdotação

Validez de las pruebas de razonamiento de la batería de evaluación de altas habilidades/superdotación

Tatiana de Cássia Nakano*
Pontifícia Universidade Católica de Campinas - PUC-Campinas, Campinas, São Paulo, Brasil

Luísa Bastos Gomes**
Pontifícia Universidade Católica de Campinas - PUC-Campinas, Campinas, São Paulo, Brasil

Karina da Silva Oliveira***
Pontifícia Universidade Católica de Campinas - PUC-Campinas, Campinas, São Paulo, Brasil

Evandro Morais Peixoto****
Pontifícia Universidade Católica de Campinas - PUC-Campinas, Campinas, São Paulo, Brasil

ABSTRACT
Considering the importance of validity evidence studies when designing psychological instruments, this study intended on verifying the construct validity of reasoning subtests, through convergent evidence and CFA, as part of a battery for giftedness assessment. A Manova was also applied to evaluate sex and grade differences. The sample was composed of 96 students, in which 49 were female, with 6th grade (n=15), 7th grade (n=12), 8th grade (n=49) students and 2nd year (n=20) students from a public school in the state of São Paulo, aged between 10 and 18 (M = 13.4 years, SD = 1.8). The participants answered the Raven Progressive Matrices Test (general scale) and reasoning subtests that belong to the Battery for Giftedness Assessment (BAHA/G). The results from Pearson's correlation indicated convergence as the majority of the factors composed in BAHA/G showed positive and significant correlations with the Raven test factors, and the CFA displayed two latent variables with strong correlations, particularly among their total score (r=.976). With these results, we found the moderate relationship between these instruments. Further studies are recommended on other types of validity evidence of these subtests to confirm its psychometric qualities.

Tatiana de Cássia Nakano, Luísa Bastos Gomes, Karina da Silva Oliveira, Evandro Morais Peixoto

RESUMO

Considerando a importância da busca por evidências de validade no processo de construção de instrumentos psicológicos, este estudo teve como objetivo verificar a validade de construto dos subtestes de raciocínio, da bateria de avaliação das altas habilidades / superdotação (BAAH/S), através de evidências de convergência e análise fatorial confirmatória. Também foi realizada a análise multivariada da variância para avaliar a influência das variáveis sexo e série. A amostra foi composta por 96 estudantes, 49 do sexo feminino, do 6º ano (n=15), 7° ano (n=12), 8° ano (n=49) do ensino fundamental II e 2º ano (n=20) do ensino médio de uma escola pública do Estado de São Paulo, com idades entre 10 e 18 anos (M=13,4 anos; DP=1,8). Os participantes responderam ao teste Matrizes Progressivas de Raven (escala geral) e aos Subtestes de Raciocínio da Bateria de Altas Habilidades e Superdotação (BAAH/S). Os resultados da correlação de Pearson indicaram convergência entre as medidas dos fatores da BAAH/S e correlações positivas com os fatores do teste do Raven. A AFC demonstrou duas variáveis latentes com fortes correlações, principalmente considerando o escore total (r=0,976) Tais resultados demonstram uma relação moderada entre os instrumentos em questão. Novos estudos voltados à investigação de outros tipos de evidências de validade são recomendados.

Palavras-chave: evidências de validade, altas habilidades, superdotação, inteligência, Matrizes Progressivas de Raven.

RESUMEN

Teniendo en cuenta la importancia de evidencias de validez en el proceso de construcción de instrumentos psicológicos, este estudio tuvo como objetivo verificar la validez de criterio de las pruebas de razonamiento de la Batería de Evaluación de altas habilidades/superdotación (BAAH/S). También fue realizado Manova para evaluar las diferencias entre sexo y ciclo. La muestra fue compuesta por 96 alumnos, 49 mujeres, el 6º año (n=15), 7 años (n=12), 8 año (n=49) de la escuela primaria y 2º año (n=20) de la escuela secundaria en una escuela pública del estado de São Paulo, con edades comprendidas entre los 10 y 18 años (M=13,4 años, SD=1,8). Los participantes respondieron al Test de matrices Progresivas de RAVEN - Escala General y la (BAAH/S). Los resultados de correlación de Pearson indicaron convergencia entre los factores de la BAAH/S y correlaciones positivas con los factores del test de RAVEN, y la AFC demostró dos variables latentes con correlaciones fuertes, teniendo en cuenta la puntuación total (r =. 976) estos resultados demuestran una relación moderada entre los instrumentos, lo que confirma la validez criterio de la prueba en el desarrollo. Además, se recomienda nuevos estudios de otros tipos de evidencia de validez.

Palavras-chave: evidencias de validez, altas habilidades, superdotación, inteligencia, Matrices Progresivas de Raven.

Tatiana de Cássia Nakano, Luísa Bastos Gomes, Karina da Silva Oliveira,
Evandro Morais Peixoto

For years remembered in history, men and women with high skills led important advances in different fields of knowledge, becoming a subject of interest for society. Given the importance of this phenomenon, researchers have focused their efforts on three main objectives: (1) identifying variables that predict high-performance, (2) how to use these variables in intervention programs and (3) how to evaluate the effectiveness of these programs, as methodological challenges were found when researching on giftedness (Subotnik, Olszewski-Kubilius, & Worrell, 2011).

For a long time, psychologists – mainly in the area of psychometrics - associated giftedness with intelligence, which was measured by high scores on I.Q. tests (Guilford, 1967; Kaufman & Sternberg, 2008; Reis & Renzulli, 2009). This prospect has changed, as a large number of researchers started pointing out the limitations on the use of intelligence tests concerning on how to define and identify a person with high skills (Gardner, 1983; Sternberg, 2005). Accordingly, many researchers have reached a consensus on the need to assume wider perspectives on intelligence as well as the use of new assessment tools in order to seize the broad spectrum of high capacity (Hérnandez-Torrano, Férrandiz, Ferrando, Prieto & Fernández, 2014). The literature has increasingly recognized that this type of instrument, used exclusively in the identification process, may not capture other areas of excellence, such as leadership or arts, so that such talents ended up lost during the process based only on general intellectual ability (Pfeiffer & Blei, 2008; Pierson, Kilmer, Rothlisberg, & McIntosh, 2012).

As a result, the definitions began changing into a more comprehensive and open approach, including several attributes besides cognitive ability, such as creativity, leadership, personal motivation, personality traits, artistic and musical ability, emotional processes and social context as components of giftedness, which then was viewed as a multidimensional phenomenon (Feldman, 2000; Gagné, 2005). Such comprehension can be found in various models (Nakano & Siqueira, 2012), for instance the Triarchic Theory of Intelligence (Sternberg, 1991), the Three Rings Conception of Giftedness (Renzulli, 1986), the Differentiated Model of Giftedness and Talent (Gagné, 2000), the Theory of Multiple Intelligences (Gardner, 1983) and the Multifactorial Model of Giftedness (Mönks, 1992).

Brazilian public policies recognize this multidimensional approach on high skills/giftedness as part of the Brazilian Special Education Program, which defines students with high abilities as those who present high potential, combined or isolated, in intellectual, academic areas, leadership and psychomotricity, in addition to high creativity, increased involvement in learning and achieving tasks regarding their interest (Brasil, 2010).

Tatiana de Cássia Nakano, Luísa Bastos Gomes, Karina da Silva Oliveira,
Evandro Morais Peixoto

Researchers have recommended the use of multiple criteria when identifying this phenomenon (such as assessing cognitive skills, creativity, motivation, leadership) given its multidimensionality, as well as the use of multiple resources as source of information (testing, portfolios, parents and teacher's indication and self-assessment), considering it the best method when identifying giftedness (Baer & Kaufman, 2005), However, note that although there are several federal laws designed to find these children which demand their identification, the overall lack of specific instruments for this population is still present (Ribeiro, Nakano & Primi, 2014). There is, so far, no instrument approved for professional use by the Brazilian Federal Council of Psychology (CFP), meaning that there are non-validated instruments for specific use, along with the still standing emphasis on cognitive assessment.

In view of this scenario, a group of researchers began the process of creating a battery to assess this phenomenon in children and adolescents (Nakano & Primi, 2012). This instrument contains reasoning subtests (verbal, numerical, logic and abstract), as well as creative verbal and figural subtest, in addition to a teacher's scale, and it was developed from already available instruments for professional use in Brazil [Battery of Reasoning Tests - Primi & Almeida, 2000; Test of Children's Figural Creativity – Nakano, Wechsler & Primi, 2011], as well as other tests that show validity evidence [Metaphors Creation Test - Primi, 2014]. Various studies have been conducted with the battery, investigating correlations between the intelligence subtests and creativity, as well as factorial structure through confirmatory analysis (Nakano, Wechsler, Campos & Milian, 2015) and searching for evidence of validity of criterion (Nakano, Primi & Ribeiro, 2014). The results were positive, considering the study of the exploratory factor analysis indicated three factors explaining 70.72% of the total variance, and the confirmatory factor analysis study indicated that intelligence, figural and verbal creativity from the battery are independent factors when compared to another creativity assessment instrument.

Considering the search for evidence of validity is presented as one of the most important pointers in the process of building an instrument (Nunes & Primi, 2010), the literature has recommended sources of information which usually include evidence of concurrent validity, using other similar instruments to measure specific constructs (Peters & Gentry, 2012), as focused on the study presented here. In this case, it is expected moderate correlations, ranging from .20 to .50 (Primi, Muniz & Nunes, 2009).

The ongoing course of complementing the battery with studies of this nature made way for others, seen in a comparison study of its figural creativity subtest results with the Brazilian Child and Adolescent Figural Creativity Test (Gomes & Nakano, 2014) and, in another

study, with the test Thinking Creatively with Figures in Torrance (Abreu & Nakano, 2014), and the verbal subtest was compared to the test's results of Thinking Creatively with Words in Torrance (Miliani & Nakano, 2014). All showed evidence of convergent validity, which stimulated the gathering of the same kind of research on intelligence subtests.

The choice of instrument taken as a criteria in the search of validity evidence was made considering it has been widely used in the scientific literature as measure when evaluating high cognitive skills/giftedness (Leikin, Paz-Barich & Leikin, 2013; Lohman, Korb & Lakin, 2008; Mills, Ablard & Brody,1993; Van der Ven & Ellis, 2000; Shaunessy, Karnes & Cobb, 2004; Yakmaci-Guzel & Akarsu, 2006). Added to this is the fact that the Raven's progressive matrices, the Stanford-Binet and Wechsler scales were cited in the literature as the most used when identifying gifted children around the world, as well as considered a criteria in the processes of construction and adaptation of new instruments (Silverman, 2009).

Considering the search for evidence of validity is presented as one of the most important pointers in the process of building an instrument (Nunes & Primi, 2010), the literature has recommended sources of information which usually include evidence of convergent validity, using other similar instruments to measure specific constructs. This way, convergent validity represents the convergence between different methods of assessing similar constructs (Peters & Gentry, 2012).

When evaluating this type of validity, usually researchers use correlations between the raw scores presented by the participant in two or more instruments referring to the same construct. However, important criticisms have been presented regarding this procedure, once this correlation is not need due to measurement error associated with the scores of each instrument (Courvoisier & Etter, 2008). According to these authors, this problem can be addressed by the use of latent variable models, which separate measurement error from true individual differences. Thus, the correlations obtained between true scores are better estimates of the true relations between constructs. Considering this perspective, the Confirmatory factor analysis (CFA) was employed to estimate the correlations between the latent variables projected through the chosen instruments (BAAH and Raven).

Method

Participants

The sample was composed of 96 students, 47 males (49%), from the 6th grade (n = 15), 7th grade (n = 12), 8th grade (n = 49) elementary school and 2nd year (n = 20) high school from a public school in the State of São Paulo, selected by convenience. Participants aged from 10 to 18 years old (M = 13.4 years; DP = 1.8).

Instruments

Battery for the Assessment of High Abilities/Giftedness (Nakano & Primi, 2012).
The test is composed of subtests involving the assessment of intelligence and creativity constructs (figurative and verbal forms) and a teacher's assessment scale. In this study, only those subtests concerning the evaluation of reasoning were used.
The items related to the assessment of intelligence include various types of reasoning (verbal, numerical, abstract and logic). The conception of these items was based on the arrangements and categories of already existing items from an instrument in use and approved by the Brazilian CFP (Bateria de Provas de Raciocínio – BPR-5, Almeida & Primi, 2000), such as its child version (Bateria de Provas de Raciocínio Infantil – BPRi, Primi & Almeida, not published), validated for use in Brazil. This test is based on the possibility of simultaneous evaluation of the g factor and more specific factors (Primi & Almeida, 2000) by means of reasoning inductive-type problems.
Raven's Progressive Matrices – General Scale (Campos, 2008)
The test has been considered a classic instrument for evaluating aspects related to the intellectual potential of a wide age range and educational level, becoming suitable for evaluating teenagers – ranging from 11 years old to adults (Cunha, 2000). It was drawn based on the bifatorial theory of Charles Spearman and aims to evaluate the general intellect-"g" factor, more specifically one of its components, titled "edutive" capability, in other words, the ability to recognize relations. (Bandeira, Alves, Giacomel & Lorenzatto, 2004).
Regarding the psychometric properties of this instrument, Gonçalves and Fleith (2011) claim that there are sufficient studies favoring the use of the test in national territory. It is important to note that the Brazilian Federal Council of Psychology approves this instrument through its evaluation system of psychological tests (SATEPSI).
Procedure

Tatiana de Cássia Nakano, Luísa Bastos Gomes, Karina da Silva Oliveira,
Evandro Morais Peixoto

Initially this project was submitted to the Research Ethics Committee from PUC- CAMPINAS, under the protocol CAAE: 21487513.7.0000.5481, which approved the conduct of this research. Then, the terms and consents were sent to the parents. Students whose parents or guardians signed the term approving the involvement of the child in accordance completed the tests mutually in the classroom, with the approximate period of one hour and thirty minutes.

The CFA was conducted through the statistical program Mplus 7.3 (Muthén & Muthén, 2012). Specifically, we have applied the Robust maximmum Likelihood estimation. The choice of this procedure was based on the appropriateness of these methods for asymmetrical variables, not normal. The model was tested from the recommended adjustment indices by Schweizer (2010): chi-squared (χ^2), degrees of freedom (df), reason regarding degrees of freedom (χ^2/df), *Comparative Fit Index* (CFI)*, Root-Mean-Square Error of Approximation* (RMSEA) e *Standardized Root Mean Square Residual* (SMRS). As for the reference values, we have adopted those commonly employed in literature: $\chi^2/gl \leq 3$; CFI $\geq 0,95$; RMSEA $\leq 0,05$ e SMRS $\leq 0,08$. We then used the multivariate analysis of variance (Manova) to check the influence of sex and grade, considering the hypothesis that gender does not influence the participants' performance, although school grade should demonstrate influence over the variables regarding intelligence subtests.

Bearing in mind the scientific literature's recommendations on the importance of finding correlations above 0.50 between the constructs measured in this type of validity evidence (Nunes & Primi, 2010), such criteria was regarded as an expected result in this work.

Results

Initially the calculated descriptive statistics for each measure considered both the four types of reasoning from the subtests of the BAHA/G as well as those obtained in the five series of Raven, separated according to the participant's gender. Considering the apparent average difference between the participants' sexes, the univariate analysis of variance was performed.

Table 1

Participants' Means and Standard Deviations in both instruments (RAVEN & BAHA/G) regarding the variable Sex

		Female		Male		Total	
		M	*SD*	*M*	*SD*	*M*	*SD*
RAVEN	Series A	10.33	1.62	10.47	1.38	10.40	1.50
	Series B	9.76	2.64	9.00	2.81	9.39	2.73
	Series C	7.12	2.80	7.00	2.92	7.06	2.85
	Series D	5.88	3.83	5.28	3.58	5.58	3.70
	Series E	1.59	2.15	1.60	1.70	1.59	1.96
	Total	34.67	9.52	33.34	9.79	34.02	9.63
BAHA/G	VR	6.24	1.40	5.81	2.12	6.03	1.79
	AR	7.02	2.37	6.15	2.88	6.59	2.66
	NR	5.96	3.35	2.91	3.24	5.94	3.28
	LR	18.94	7.35	17.19	9.20	18.08	8.31
	Total	38.16	35.06	35.06	13.84	36.65	12.61

Note. VR = verbal reasoning, AR = abstract reasoning, NR = numerical reasoning, LG = logical reasoning.

The results indicated that this variable showed significant influence, in the Raven test, on the results of the B-series (F = 9.13; p=0.003) and total result (F = 4.44; p=0.38). Concerning the BAHA/S subtests, they proved to be significant as seen in the results obtained in the abstract reasoning subtest (F = 5.93; p=0.17), logical reasoning (F = 3.95; p=0.05) and for the total of the instrument (F = 5.10; p=0.02). This way it is possible to affirm that the female participants showed superior results when are compared to male participants in all measures in which this variable proved to be significant. However, we recommend caution when interpreting this result, since perceived differences may be due to real differences in the rated construct or on the items' parameters. Considering that the same analysis via Item Response Theory was not performed, parameter invariance related issues due to the differential functioning of items should not be discarded and deserve to be investigated in further studies. Formerly, the descriptive statistics of the sample, considering the participants' grades was held. The results are displayed on Table 2.

Table 2
Participants' Means and Standard Deviations in both instruments (RAVEN & BAAH/S) regarding Grades

		6th Grade		7th Grade		8th Grade		2nd Year	
		M	SD	M	SD	M	SD	M	SD
RAVEN	Series A	9.47	1.68	10.25	1.42	10.59	1.55	10.70	1.03
	Series B	7.33	4.99	9.00	2.62	9.69	1.88	10.40	1.18
	Series C	4.60	3.75	5.83	2.79	7.59	2.21	8.35	2.25
	Series D	0.73	1.83	5.42	3.28	7.16	2.72	5.45	4.01
	Series E	0.20	0.77	1.42	1.50	2.29	2.16	1.05	1.63
	Total	22.33	9.90	31.92	9.37	37.33	8.19	35.95	5.31
BAHA/G	VR	6.00	1.30	5.00	2.08	5.92	1.73	6.95	1.79
	AR	5.60	2.38	5.42	2.67	6.84	2.71	7.45	2.41
	NR	4.33	2.96	5.08	3.02	6.51	3.47	6.25	2.86
	LR	9.47	8.04	16.42	8.82	19.90	7.36	21.10	6.12
	Total	25.40	12.43	31.92	14.22	39.16	11.40	41.75	8.84

Considering the means, a predisposition to obtain better results was found as the grades progressed, which was expected according to the intelligence theories. In the Raven test, this result happens in series A, B and C and in VR, AR, LR subtests and total score from BAHA/G. Surprisingly, students from 8th grade presented a higher average then students from 2nd year in series D, E and total score from Raven and subtest RN from BAHA/G.

Once again, considering the apparent differences in average, the univariate analysis of variance was used to investigate the influence of the variable grade. The results indicated significant influence in all series of RAVEN instrument, namely, series A (F = 3.72; p=0.014), series B (F = 7.74; p=0.0001), series C (F = 7.38; p=0.0001), series D (F = 15.35; p=0.0001), and series E (F = 4.95; p=0.003) and also for the total score of the instrument RAVEN (F = 14.42; p=0.0001). The variable also presented significant influence on AR subtests (F = 2.84; p=0.042) and LR (F = 10.46; p=0.0001) of the BAHA/G, as well as for the total of this instrument (F = 8.18; p=0.0001). Regarding the interaction of the variables sex and grades, the only influence found was on series B from RAVEN (F = 4.50; p=0.005).

Then, the CFA was used in order to verify the main objective of this study: the relationship between the two latent variables estimated by the instruments. Bearing in mind the theoretical proposal to be tested, the model was organized in order to contemplate two latent variables. The first, named Eductive Capacity, which reflected the results presented by the students in the five series from the Raven test, and the second, called Inductive Capacity, which reflected the four types of reasoning assessed by BAAH. The initial results indicated levels of adjustments considered poor: $\chi^2 = 54.920$, df =

26, $p < 0,01$, $\chi^2/df = 8,042$, $CFI = 0,904$; $RMSEA = 0,108$ (I.C. 90%= 0,068-0.147); $SMRS = 0.06$. However, the modification indexes indicated a correlation between the variables series "D" and "E", whereas both variables evaluate the most difficult plots of a same construct. Therefore, we had opted to insert a parameter that was not previously considered. After the insertion of the parameter, the results presented adjustment indexes which were considered good: $\chi^2 = 36,255$, $df = 25$, $p = 0,11$, $\chi^2/df = $, $CFI = 0,971$; $RMSEA = 0,068$ (I.C. 90%= 0,00-0,108); $SMRS = 0.04$. Briefly, the model is displayed on Image 1, where the factorial charges presented by the observed variables and correlations index between latent variables.
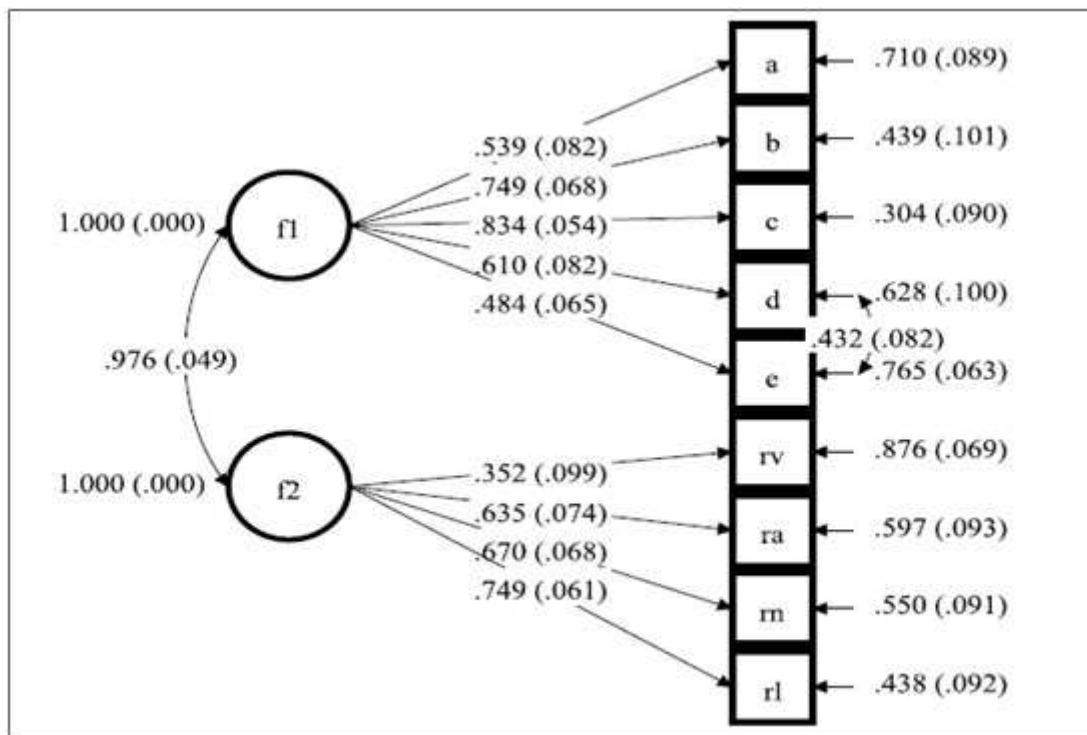

Image 1: Confirmatory Factorial Analysis – Factorial Charges between instruments

The results displayed validity evidence between both measures, whereas the correlation among the factors estimated from BAHA/G and Raven presented significant value (r=.976). In addition, it appears that the charges made by the factors in the respective variables are greater than 0,352 (RL). A more detailed analysis of the correlations between the variables can be seen on Table 3, it was estimated for both isolated measures of each test.

Table 3
Correlations between subtests

	RAVEN Series A	RAVEN Series B	RAVEN Series C	RAVEN Series D	RAVEN Series E
BAHA/G VR	.085	.179	.335**	.083	.221*
BAHA/G AR	.427**	.461**	.489**	.346**	.284**
BAHA/G NR	.331**	.495**	.565**	.408**	.331**
BAHA/G LR	.370**	.578**	.562**	.578**	.370**

*$p \leq .05$; **$p \leq .01$

Significant positive correlations were found between the AR subtests, NR, LR and BAHA/G's total score with all series and Raven's total score, varying between r = .221 and r = .633. The only exception occurred on subtest VR, whose correlation with series A, B and D of the Raven were not significant, even with the other (r = .335 with Series C and r = .221) and series E and the total (r = .240) values have been significant.

While executing this procedure, no important alterations were verified within the correlations, considering the fact none of the items sets showed very high internal consistency. In fact, Raven's values oscillated between 0,60 and 0,86; whereas the BAAH/G fluctuated between 0,83 and 0,87. This way, the authors chose to keep the information on Table 1. It is worth mentioning this procedure was performed following the recommendations of the scientific literature (Schumacker & Muchinsky, 1996).

Discussion

Based on the extracted information from this analysis, we found that the present study showed statistically significant correlations compatible with the equivalent construct assessing instrument, confirming its adequate convergence and specificity. While the literature explains the concept of validity as essential during the development and evaluation of tests, based on the obtained scores and performance, we can affirm the validity of reasoning subtests from the battery (American Psychological Association, 1999).

It is important to stress that positive and significant correlations between the instruments were expected since both are based on a theory which considers a general intellectual ability, called "g" factor (Almeida & Primi, 2000). Though differences might be noted concerning the intelligence model that supports the instruments, with Raven marked by a general factor but with tasks which involve

eductive reasoning, BAHA/G's activities reflect, most directly, inductive reasoning, evaluated by means of items involving fluid and crystallized intelligence on specific skills.

The first type could be understood as the individual's ability to perform mental operations when facing a new task that cannot be performed automatically, associated with little prior knowledge components and cultural influence, depending more on biological and genetic factors (Almeida, Lemos, Guisande, & Primi, 2008). On the other hand, the second one would be developed from cultural and educational experiences, found in most school activities, in order to represent different aspects of required capabilities in solving everyday problems, and can be understood as a form of social intelligence or common sense, in which intellectual investment is required when learning (Campos & Nakano, 2012). In this sense, future research on BAHA/G could set a model with a g factor explain those correlations between BAHA and Ravens scores.

A hypothesis raised on the lower correlation found in subtest VR with Raven's total score, as well as the absence of significant correlation with two of its series, suggest findings that the VR (verbal reasoning) is the only factor of BAHA/G that evaluates both the Fluid Intelligence as well as the Crystallized. Thus why it may not have found significant survey directories, once the proposal of Raven's progressive matrices would focus on assessing only the General Intelligence (Almeida, 2009). Similarly, the VR in this battery specifically assesses the extent of vocabulary and the ability to establish abstract relations among verbal concepts, involving fluid and crystallized intelligence, the second type is not included in the Raven test, seen as the examinee must only complete figures, selecting, among the alternatives, one that would complete the series. Although scientific literature has been showing evidence on the association between fluid and crystallized intelligence, expecting, in this case, some relation between such measures, the result was not found in this study. A hypothesis to consider might also be explained by the limited statistical power due the sample size.

Note the fact that, in his introduction to matrices, Raven argues that the test alone does not evaluate the General Intelligence, suggesting instead the combination of a vocabulary test to supplement the Matrices, as well the use of language (Silverman, 2009). Thus, the importance of also creating a subtest taking this aspect into account is shown, as found in the VR BAHA/G subtest, and that gap can also justify the lower correlations found in this subtest and the instrument taken as a criterion. In a similar study, when comparing the advanced Raven with the abstract reasoning of BPR-5, which gave rise to the BAHA/G subtests, indicated 0.48 correlation between measures (Nunes et al., 2012), confirming the results found.

Another important matter refers to the influence of sex and grade on cognitive performance. The results indicated significant differences regarding the performance of the participants by number, considering higher means obtained by students in more advanced series, with a few exceptions, and the data acquired from the scientific literature. Almeida, Lemos, Guisande and Primi (2008), for example, stated that several authors refer to a differential impact of schooling in different cognitive skills, including those related to fluid intelligence. The authors, when researching a sample of Portuguese students, also reported results of evolution in cognitive achievement (in verbal, numerical reasoning subtests, abstract, mechanical and spatial reasoning tests, which formed the basis for the evidence that comprises the BAHA/G) as the school ages advance. Likewise, several other studies (Nakano, 2012; Wechsler et al., 2010a) verified educational series impact in the development of intellectual skills. These results agree with those reported in this study.

Regarding the influence of the variable sex in cognitive performance, the results of this study indicate its significance in some measures of both instruments, favoring females. Hallinger and Murphy (1986) suggest that sex can interfere with academic performance, in which girls display an anxious behavior in early academic life compared to boys and as a result, these girls end up studying more and getting better results than the opposite sex.

The variance of performance in this construct, concerning sex, is a well-known controversial issue. While authors support the idea that differences between men and women are to be expected in some of the intellectual skills, but not in the total results in intelligence (Wechsler et al., 2010b), other studies point the absence of this difference (Nakano, 2012). Likewise, opposing results are also reported, with better male performance on tasks of vocabulary and crystallized intelligence (Wechsler, Vendramini & Schelini, 2007).

According to the results found in several studies regarding the sex variable, the differences between the sexes could be related to cognitive skills and are not necessarily specific to the g factor, indicating equivalent performance between genders (Rueda & Castro, 2013; Keith, Reynolds, Patel & Ridley, 2008; Sluis et al., 2008; Weiss, Kemmler, Deisenhammer, Fleischhacker, & Delazer, 2003). This way, men would have a better performance in visuospatial skills, mathematical reasoning and mechanics, while women would have a better performance in tasks involving verbal skills and perceptual speed (Rueda & Castro, 2012).

Final Considerations

Bearing in mind the goal of this study was a search for evidence of validity in reasoning subtests of the Battery for the Assessment of High Abilities and Giftedness, it was possible to verify that the analysis of the results pointed to significant moderate correlations between most of the features of both instruments, as well as among the most general factors.

The research illustrates its importance when it comes to ways of identifying giftedness, as the perception obtained by national scientific literature emphasizes the need for developing instruments targeting this specific population, and the implications resulting from the lack of studies in this area. Highlights are needed on other studies involving this battery under study, so that its psychometric qualities can be confirmed.

References

Abreu, I., & Nakano, T. (2014). Correlação entre Testes de Criatividade Figural [Correlation between Figural Creativity Tests]. In XIX Encontro de Iniciação Científica. Campinas: PUC-Campinas.

Almeida, F. (2009). Teste das Matrizes Progressivas de Raven [Raven's Progressive Matrices]. *Peritia: Revista Portuguesa de Psicologia*. Retirado de http://www.revistaperitia.org/wpcontent/uploads/2010/04/MPCR.pdf. Acesso em 15/06/2015.

Almeida, L.S., Lemos, G., Guisande, M.A., & Primi, R. (2008). Inteligência, escolarização e idade: normas por idade ou série escolar? [Intelligence, education and age: standards by age or school grade?] . *Avaliação Psicológica*, 7(2), 117-125.

Almeida, L.S., & Primi, R. (2000). *Baterias de Prova de Raciocínio – BPR-5* [Reasoning Battery Tests] São Paulo: Casa do Psicólogo.

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

Baer, J., & Kaufman, J. C. (2005). Bridging Generality and Specificity: The Amusement Park Theoretical (APT) Model of Creativity. *Roeper Review,* 27, 158-163.

Bandeira, D. R., Alves, I. C. B., Giacomel, A. E., & Lorenzatto, L. (2004). Matrizes progressivas coloridas de Raven – escala especial: normas para Porto Alegre/RS [The Raven's coloured

progressive matrices: norms for Porto Alegre, RS]. *Psicologia em Estudo*, 9(3), 479-486.

Brasil. (2010). *Educação de Alunos Superdotados/Altas Habilidades* [Education on Gifted/High Abilities Students]. Brasília: Ministério da Educação e do Desporto / Secretaria de Educação Especial.

Campos, F. (2008). *Matrizes progressivas de Raven – escala geral* [Raven's Progressive Matrices – General Scale]. São Paulo: Cepa.

Campos, C. R., & Nakano, T. C. (2012). Produção científica sobre avaliação da inteligência: o estado da arte [Scientific production on intelligence assessment: the State of the art]. *Interação em Psicologia*, 16(2), 271-282.

Courvoisier, D., & Etter, J. F (2008). Using item response theory to study the convergent and discriminant validity of three questionnaires measuring cigarette dependence. *Psychology Addict Behavior,*22(3), 391-401.

Cunha, J. A. (2000). Módulo VII – catálogo de técnicas úteis [Module VII-useful technical catalog]. In Cunha, J. A., *Psicodiagnóstico – V*. Porto Alegre: Artmed.

Feldman, D. H. (2000) The development of creativity. In R. J. Sternberg (Ed), *Handbook of creativity* (pp. 169-189). Cambridge: Cambridge University Press.

Gagné, F. (2000). Understanding the complex choreography of talent development through DMGT-Based Analysis. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent (*pp.67-79). 2. ed. Oxford: Pergamon.

Gagné, F. (2005). From gifted to talents: the DMGT as a developmental model. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of Giftdeness* (pp. 98-120). New York: Cambridge University Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York, NY: Basic Books.

Gomes, L. B., & Nakano, T. (2014). Validade Convergente do Subteste de Criatividade Figural da Bateria de Avaliação das Altas Habilidades/Superdotação [Convergent Validity in Figural Creativity Subtest from a Battery of Assessment of High Abilities and Giftedness]. In XIX Encontro de Iniciação Científica. Campinas: PUC-CAMPINAS.

Gonçalves, F. C., & Fleith, D. S. (2011). Estudo comparativo entre alunos superdotados e não-superdotados em relação à inteligência e criatividade [Comparative study between gifted and non-gifted students related to intelligence and creativity]. *Psico*, 42 (2), 263–268.

Guilford, J. P. (1967). *The nature of human intelligence.* New York, NY: McGaw-Hill.

Hallinger, P., & Murphy, J. F (1986). The social context of effective schools. *American Journal of Education, 94*(3), 328-355.

Hérnandez-Torrano, D., Férrandiz, C., Ferrando, M., Prieto, M. D., & Fernández, M. C. (2014). La teoria de las inteligências múltiples en la identificación de alumnos de altas habilidades (superdotación y talento) [The Theory of Multiple Intelligences in the Identification of Students with High Abilities (Giftedness and Talent)]. *Anales de Psicologia, 30*(1), 192-200.

Kaufman, S. B., & Sternberg, R. J. (2008). Conceptions of giftedness. In S. Pfeiffer (Ed.), *Handbook of giftedness in children: Psycho-Educational theory, research and best practices* (pp. 71-91). New York: Springer.

Keith, T. Z., Reynolds, M. R. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: evidence from the Woodcock-Johnson III tests of cognitive abilities. *Intelligence*, *36*(6), 502-525.

Leikin, M., Paz-Baruch, N., & Leikin, R. (2013). Memory Abilities in Generally Gifted and Excelling-in-mathematics Adolescents. *Intelligence*, 41, 566–578.

Lohman, D. F., Korb, K., & Lakin, J. (2008). Identifying academically gifted English language learners using nonverbal tests: A comparison of the Raven, NNAT, and CogAT. G*ifted Child Quarterly*, *52*, 275-296.

Miliani, A. F. M., & Nakano, T. C. (2013). *Bateria de avaliação das altas habilidades / superdotação: construção e estudos psicométricos* [Battery for Assessment of High Abilities and Giftedness: Construction and Psychometric Studies]. In XVII Encontro de Iniciação Científica. Campinas: PUC-CAMPINAS.

Mills, C. J., Ablard, K. E., & Brody, L. E. (1993). The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students. *Roeper Review*, *15*(3), 183-186.

Mönks, F. J. (1992) Development of gifted children: The issue of identification and programming. In F. J. Mönks & W. Peters (Eds.), *Talent for the future* (pp. 191-202)*.* Assen: Van Gorcum.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus User's Guide* (6a ed.). Los Angeles, CA: Muthén & Muthén.

Nakano, T. C. (2012). Criatividade e inteligência em crianças: habilidades relacionadas [Creativity and intelligence in children: related skills]. *Psicologia: Teoria e Pesquisa,* 28(2), 145-159.

Nakano, T. C., & Primi, R. (2012). *Bateria de Avaliação das Altas Habilidades/Superdotação* [Battery for Assessment of High Abilities/Giftedness]. Not published.

Nakano, T. C., Primi, R., & Ribeiro, W. J. (2014) Validade da estrutura fatorial de uma bateria para avaliação das altas habilidades / superdotação [Validity of the Factor Structure of a Battery of Assessment of High Abilities]. *Psico, 45*, 100-109.

Nakano, T. C., & Siqueira, L. G. G. (2012). Superdotação: modelos teóricos e avaliação [Giftedness: theoretical models and evaluation]. In C.S. Hutz (Org.), *Avanços em Avaliação Psicológica e Neuropsicológica de crianças e adolescentes II* (pp.277-311). São Paulo: Casa do Psicólogo.

Nakano, T. C., Wechsler, S. M., Campos, C. R., & Milian, Q. G. (2015). Intelligence and Creativity: Relationships and their Implications for Positive Psychology. Psico-USF, 20, 195-206.

Nakano, T. C., Wechsler, S. M., & Primi, R. (2011). *Teste de Criatividade Figural Infantil: Manual técnico [Test of Figural Creativity for Children: Manual].* São Paulo: Editora Vetor.

Nunes. M. F. O., Borosa, J. C., Nunes, C. H. S. S., & Barbosa, A. A. G. E. (2012). Avaliação da Personalidade em Crianças e Adolescentes: Possibilidades no Contexto Brasileiro [Personality Assessment in Children and Adolescents: Possibilities in a Brazilian Context]. In: *Avanços em Avaliação Psicológica e Neuropsicológica de Crianças e Adolescentes II*. São Paulo: Casa do Psicólogo, 1, 253-276.

Nunes, C. H. S. S., & Primi, R. (2010). Aspectos técnicos e conceituais da ficha de avaliação dos testes psicológicos [Technical and conceptual aspects of the assessment form of psychological testing]. In Conselho Federal de Psicologia, (Org.), *Avaliação Psicológica: Diretrizes na regulamentação da profissão* (pp. 101-128) Brasília: CFP.

Peters, S. J., & Gentry, M. (2012) Internal and external validity evidence of the HOPE Teacher Rating Scale. *Journal of Advanced Academics, 23*(2), 125-144.

Pfeiffer, S. I., & Blei, S. (2008). Gifted identification beyond the IQ test: Rating scales and other assessment procedures. In S. I. Pfeiffer (Org.), *Handbook of giftedness in children* (pp. 177-198). NY: Springer Science+Business Media, LLC.

Pierson, E. E., Kilmer, L. M. Rhlisberg, B. A., & McIntosh, D. E. (2012). Use of brief intelligence tests in the identification of giftedness. *Journal of Psychoeducational Assessment, 30*(1), 10-24.

Primi, R. (2014). Divergent productions of metaphors: Combining many-facet Rasch measurement and cognitive psychology in the assessment of creativity. *Psychology of Aesthetics, Creativity, and the Arts, 8*(4), 461-474.

Primi, R., & Almeida, L. S. (2000). Estudo de validação da Bateria de Provas de Raciocínio (BPR-5) [Evidence Validation Study of the

Reasoning Battery]. *Psicologia: Teoria e Pesquisa*, 16(2), 165-173.

Primi, R., Muniz, M., & Nunes, C. H. S. S. (2009). Definições contemporâneas de validade de testes psicológicos. In C. S. Hutz (Org.), *Avanços e polêmicas em avaliação psicológica* (pp. 243-266). São Paulo: Casa do Psicólogo.

Reis, S. M., & Renzulli J. S. (2009). Myth 1: The gifted and talented constitute one single homogeneous group and giftedness is a way of being that stays in the person over time and experiences. *Gifted Child Quarterly*, 53, 233–235.

Renzulli, J. S. (1986). The three-ring conception of giftdness: A developmental model for creative productivity. In R. J. Sternberg & J. E. Davis (Eds.), *Conception of giftedness* (pp. 53-92). New York: Cambridge University Press.

Ribeiro, W. J., Nakano, T. C., & Primi, R. (2014). Validade da Estrutura Fatorial de uma Bateria de Avaliação de Altas Habilidades [Factorial Structure Validity of a Battery for Assessment of High Skills]. *Psico, 45*(1), 100-109.

Rueda, F. J. M. & Castro, N. R. (2013). Análise das variáveis idade e sexo no desempenho do teste de inteligência (TI) [Analysis of the variables age and sex in the performance of intelligence test (IT)]. *Psicologia: teoria e prática*, *15*(2), 166-179.

Schumacker R. E., Muchinsky P. M. (1996) Disattenuating correlation coefficients. *Rasch Measurement Transactions*, 10(1), 479. [http://www.rasch.org/rmt/rmt101g.htm]

Shaunessy, E., Karnes, F. A., & Cobb, Y. (2004). Assessing culturally diverse potentially gifted students with nonverbal measures of intelligence. *The School Psychologist*, *58*, 99-102.

Silverman, L. K. (2009). The measurement of giftedness. In L. Shavinina (Ed.), *The international handbook on giftedness* (pp. 957–970). Amsterdam: Springer Science.

Sluis, S. V. D., Deron, C., Thiery, E., Bartels, M., Polderman, T. J. C., Verhulst, F. C., Jacobs, N., Gestel, S. V., Geus, E. J. C., Dolan, C. V., Boosmsma, D. I., & Posthuma, D. (2008). Sex differences on the Wisc-R in Belgium and The Netherlands. *Intellligence*, *36*, 48-67.

Sternberg, R. J. (1991). A triarchic view of giftedness: theory and practice. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 43-53). New York: Pearson.

Sternberg, R. J. (2005). Intelligence. In K.J. Holyoak & R.G. Morrison, *The Cambridge handbook of thinking and reasoning* (pp. 751-773). New York: Cambridge University Press.

Subotnik, R. F., Olsezwski-Kubilus, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest*, 12(1), 3-54.

Van der Ven, A. H. G. S., & Ellis, J. L. (2000). A Rasch analysis of Raven's standard progressive matrices. *Personality and Individual Differences*, 29, 45-64.

Wechsler, S. M., Nunes, M. F., Schelini, P. W., Ferreira, A. A., & Pereira, D. A. P. (2010). Criatividade e inteligência: analisando semelhanças e discrepâncias no desenvolvimento [Creativity and intelligence: analyzing developmental similarities and discrepancies]. *Estudos de Psicologia (Natal)*, *15*(3), 243-250.

Wechsler, S. M., Nunes, C. H. S., Schelini, P. W., Pasian, S. R., Homsi, S. V., Moretti, L., & Anache, A. A. (2010). Brazilian adaptation of the Woodcock-Johnson. *School Psychology International, 31*, 409-421.

Wechsler, S. M., Vendramini, C. M., & Schelini, P. W. (2007). Adaptação brasileira dos testes verbais da Bateria Woodcock-Johnson III [Brazilian Adaptation of the Verbal Tests Woodcock-Johnson III Battery]. *Interamerican Journal of Psychology, 41*, 285-294.

Weiss, E. M., Kemmler, G., Deisenhammer, E. A., Fleischhacker, W. W., & Delazer, M. (2003). Sex differences in cognitive functions. *Personality and Individual Differences*, *35*, 863-875.

Yakmaci-Guzel, B., & Akarsu, F. (2006). Comparing overexcitabilities of gifted and non-gifted 10th grade students in Turkey. *High Ability Studies, 17*(1), 43-56.

Endereço para correspondência
Tatiana de Cássia Nakano
Pontifícia Universidade Católica de Campinas
Centro de Ciências da Vida, Programa de Pós-Graduação em Psicologia
Av. John Boyd Dunlop, s/n., Jardim Ipaussurama, CEP 13060-904, Campinas – SP, Brasil
Endereço eletrônico: tatiananakano@hotmail.com
Luísa Bastos Gomes
Pontifícia Universidade Católica de Campinas
Centro de Ciências da Vida, Programa de Pós-Graduação em Psicologia
Av. John Boyd Dunlop, s/n., Jardim Ipaussurama, CEP 13060-904, Campinas – SP, Brasil
Endereço eletrônico: luisagomes14@gmail.com
Karina da Silva Oliveira
Pontifícia Universidade Católica de Campinas
Centro de Ciências da Vida, Programa de Pós-Graduação em Psicologia
Av. John Boyd Dunlop, s/n., Jardim Ipaussurama, CEP 13060-904, Campinas – SP, Brasil
Endereço eletrônico: karina_oliv@yahoo.com.br
Evandro Morais Peixoto
Pontifícia Universidade Católica de Campinas
Centro de Ciências da Vida, Programa de Pós-Graduação em Psicologia
Av. John Boyd Dunlop, s/n., Jardim Ipaussurama, CEP 13060-904, Campinas – SP, Brasil
Endereço eletrônico: epeixoto_6@hotmail.com

Tatiana de Cássia Nakano, Luísa Bastos Gomes, Karina da Silva Oliveira,
Evandro Morais Peixoto

Notas
* Docente do programa de pós-graduação em Psicologia da PUC-Campinas.
** Mestranda em Psicologia como Profissão e Ciência pela PUC-Campinas.
*** Doutoranda em Psicologia como Profissão e Ciência pela PUC-Campinas.
**** Doutor em Psicologia como Profissão e Ciência pela PUC-Campinas, docente do departamento de Psicologia da Universidade de Pernambuco.