REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

INCORPORATING TEST DESIGN INTO FACTORIAL VALIDITY: A BIFACTOR REANALYSIS OF THE BAFACALO FOLDING TEST

INCORPORANDO O FORMATO DO TESTE À VALIDADE FATORIAL: UMA REANÁLISE BIFATORIAL DO TESTE DE DOBRADURAS DA BAFACALO

INCORPORANDO EL DISEÑO DEL TEST A LA VALIDEZ FACTORIAL: UN REANÁLISIS BIFACTORIAL DEL TEST DE PLEGADO DE LA BAFACALO

Cristiano Mauro Assis Gomes
Pós-Doutor em Psicologia da Educação
Universidade Federal de Minas Gerais (UFMG), Brasil
E-mail: cristianomaurogomes@gmail.com

Jhonys de Araujo
Doutor em Psicologia: Cognição e Comportamento
Universidade Federal de Minas Gerais (UFMG), Brasil
E-mail: jhonys.bio@gmail.com

Abstract

Evidence-based psychological assessment requires well-calibrated evidence about the dimensional structure of tests, since every score interpretation presupposes that structure. The open science movement has established test democratization as an important agenda, and by broadening the use of tests, democratization makes well-calibrated structural evidence even more important. The Folding Test, part of the BAFACALO battery, has been freely accessible since 2021, and its factorial structure was investigated through an exploratory strategy with the estimation procedures available at the time, yielding a general factor (visualization) and three specific factors. Because the 30 items of the test are organized into six figures, these specific factors may arise from the format of the test itself. The present study reexamines the dimensional structure of the Folding Test through a confirmatory factor analysis with 292 Brazilian public high school students. Three confirmatory factor models were compared using WLSMV estimation, a procedure specific to dichotomous items: a unidimensional model, a bifactor model derived from a previous exploratory study of the test, and a bifactor model with one specific factor per figure (figure model). All models showed acceptable fit, and both bifactor models fitted slightly better than the unidimensional model. After adjustment, the two bifactor models converged to practically the same structure, with specific factors centered on items of figures 1 and 3, and could not be distinguished by fit. The figure model was retained for its theoretical pertinence (scaled CFI = .997; scaled RMSEA = .032). The visualization factor showed high reliability (McDonald's ω = .979), whereas the specific factors showed low reliability (ω ≤ .349). The Folding Test is essentially unidimensional, and its specific factors are better explained by the test format than by distinct abilities. The study complements the democratization of the Folding Test with better-calibrated evidence about its dimensional structure.
Keywords: Visualization; intelligence; test democratization; open science; factorial validity

Resumo

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

A avaliação psicológica baseada em evidências requer evidências bem calibradas sobre a estrutura dimensional dos testes, uma vez que toda interpretação de escores pressupõe essa estrutura. A democratização de testes, impulsionada pelo movimento de ciência aberta, amplia o uso dos testes e torna essas evidências ainda mais importantes. O Teste de Dobraduras, integrante da bateria BAFACALO, está livremente disponível desde 2021. Sua estrutura fatorial foi investigada por uma estratégia exploratória com os procedimentos de estimação da época, resultando em um fator geral (visualização) e três fatores específicos. Como seus 30 itens estão organizados em seis figuras, esses fatores específicos podem decorrer do formato do teste. O presente estudo reexamina essa estrutura por meio de análise fatorial confirmatória com 292 estudantes brasileiros do ensino médio público. Foram comparados três modelos: um modelo unidimensional, um modelo bifatorial derivado do estudo anterior e um modelo bifatorial com um fator específico por figura (modelo das figuras). Os modelos foram estimados por um procedimento específico para itens dicotômicos. Todos os modelos apresentaram ajuste aceitável, e ambos os modelos bifatoriais se ajustaram ligeiramente melhor que o unidimensional. Após os ajustes, os dois modelos bifatoriais convergiram para praticamente a mesma estrutura, com fatores específicos nos itens das figuras 1 e 3, e não puderam ser diferenciados pelo ajuste. O modelo das figuras foi mantido por sua pertinência teórica (CFI escalonado = 0,997; RMSEA escalonado = 0,032). O fator de visualização apresentou alta confiabilidade (ω de McDonald = 0,979), enquanto os fatores específicos apresentaram baixa confiabilidade (ω ≤ 0,349). O Teste de Dobraduras é essencialmente unidimensional, e seus fatores específicos são mais bem explicados pelo formato do teste do que por habilidades distintas. O estudo complementa a democratização do teste com evidências estruturais mais bem calibradas.
Palavras-chave: Visualização; inteligência; democratização de testes; ciência aberta; validade fatorial

Resumen

La evaluación psicológica basada en la evidencia requiere evidencias bien calibradas sobre la estructura dimensional de los tests, ya que toda interpretación de puntuaciones presupone esa estructura. La democratización de los tests, impulsada por el movimiento de ciencia abierta, amplía su uso y hace aún más importantes estas evidencias. El Test de Plegado, integrante de la batería BAFACALO, es de libre acceso desde 2021. Su estructura factorial se investigó mediante una estrategia exploratoria con los procedimientos de estimación de la época, lo que resultó en un factor general (visualización) y tres factores específicos. Como sus 30 ítems están organizados en seis figuras, estos factores específicos pueden derivar del formato del test. El presente estudio reexamina esa estructura mediante análisis factorial confirmatório con 292 estudiantes brasileños de educación secundaria pública. Se compararon tres modelos: un modelo unidimensional, un modelo bifactorial derivado del estudio anterior y un modelo bifactorial con un factor específico por figura (modelo de figuras). Los modelos se estimaron con un procedimiento específico para ítems dicotómicos. Todos los modelos mostraron ajuste aceptable, y ambos modelos bifactoriales se ajustaron ligeramente mejor que el unidimensional. Tras los ajustes, los dos modelos bifactoriales convergieron en prácticamente la misma estructura, con factores específicos en los ítems de las figuras 1 y 3, y no pudieron diferenciarse por el ajuste. Se retuvo el modelo de figuras por su pertinencia teórica (CFI escalado = 0,997; RMSEA escalado = 0,032). El factor de visualización mostró alta fiabilidad (ω de McDonald = 0,979), mientras que los factores específicos mostraron baja fiabilidad (ω ≤ 0,349). El Test de Plegado es esencialmente unidimensional, y sus factores específicos se explican mejor por el formato del test que por habilidades distintas. El estudio complementa la democratización del test con evidencias estructurales mejor calibradas.
Palabras clave: Visualización; inteligencia; democratización de tests; ciencia abierta; validez factorial

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

1. Introduction

Evidence-based practice in psychology places particular emphasis on the scientific grounding of test scores (American Educational Research Association et al., 2014; Bornstein, 2017; International Test Commission, 2017). Among the sources of validity evidence, evidence based on internal structure (American Educational Research Association et al., 2014) holds a foundational position: any score derived from a test presupposes a dimensional structure, as a single total score presupposes that all items measure the same attribute. Factorial validity is therefore not a technical detail but the basis on which score interpretation rests. Well-calibrated evidence about the dimensional structure of tests is therefore a fundamental requirement for evidence-based psychological assessment.

The global movement toward open science in psychology has amplified this requirement, driving a growing agenda of test democratization aimed at making validated psychological instruments freely and openly accessible (Nosek et al., 2015). In low- and middle-income countries, access to psychometrically grounded instruments has historically been limited by financial and institutional barriers, making this agenda particularly consequential for research quality and professional practice (Beidas et al., 2015). By broadening the use of a test, democratization also broadens the responsibility for ensuring that the evidence about its structure is well calibrated. Making an instrument freely accessible without such evidence leaves the fundamental problem unresolved.

This problem is particularly salient in the Latin American context, where ethical and regulatory frameworks increasingly demand that practitioners rely on instruments whose scores have been rigorously evaluated and are valid for their intended uses (Conselho Federal de Psicologia, 2022; Toro et al., 2025), yet the availability of instruments meeting this standard remains uneven (Elosua & Wechsler, 2025; Hernández et al., 2021). In this context, strengthening the structural evidence of instruments that are already available is a direct contribution to research and professional practice in the region.

The Battery of High-Order Cognitive Factors (BAFACALO) is a pioneering Brazilian instrument grounded in the Cattell-Horn-Carroll (CHC) model and Carroll's

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Three-Stratum model, developed at the Laboratory for Cognitive Architecture Mapping (LAICO) in 2006 to assess the general factor of intelligence, six broad cognitive abilities, and 18 narrow abilities. Systematic validity research on the BAFACALO has produced evidence of factorial validity (Gomes, 2010b, 2011, 2012; Gomes & Borges, 2009a, 2009b; Gomes, de Araújo, et al., 2014; Gomes et al., 2022; Gomes & Golino, 2015), predictive validity (Gomes, 2010a, 2011; Gomes & Golino, 2012a; Gomes, Marques, & Golino, 2014), incremental validity (Alves et al., 2012; Gomes, 2012; Gomes et al., 2022; Gomes & Golino, 2012b), and divergent and convergent validity (Gomes, 2010b, 2011, 2012; Gomes & Borges, 2009a, 2009b). In 2021, LAICO launched a test democratization agenda directed at making psychological instruments openly and freely accessible, incorporating the BAFACALO into this initiative. The tests were progressively released, first as preprints (Gomes & Nascimento, 2021a, 2021b, 2021c, 2021d, 2021e, 2021f, 2021g, 2021h, 2021i, 2021j, 2021k, 2021l, 2021m, 2021n; Gomes et al., 2021a, 2021b, 2021c, 2021d) and subsequently in peer-reviewed journals (Gomes & Araujo, 2025a, 2025b, 2025c, 2025d). Their adoption across institutions for the study, teaching, and assessment of human cognitive abilities has made the soundness of their validity evidence more important.

The factorial structure of the Folding Test, one of the 18 tests of the BAFACALO, was investigated by Gomes and Borges (2009a) through an exploratory strategy: exploratory factor analysis identified the number of specific factors, and confirmatory factor analysis (CFA) selected among the resulting solutions, yielding a general factor and three specific factors. This strategy is appropriate, but testing models specified a priori, whenever possible, tends to provide more robust evidence of factorial structure. The Folding Test comprises 30 items organized into six figures, each with five items, so the specific factors may arise from the format of the test itself. Accordingly, the present study reexamines the dimensional structure of the Folding Test through a confirmatory comparison of three models: a unidimensional model, the bifactor model derived from Gomes and Borges (2009a), and a bifactor model whose specific factors correspond to the six figures of the test (hereafter, the figure model).

The analysis by Gomes and Borges (2009a) followed the technical possibilities of the time, adopting the CFA approach described by Byrne (2001), conducted in

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

AMOS with maximum likelihood estimation. Since then, specific estimation procedures for dichotomous items, such as weighted least squares estimation based on tetrachoric correlations, have become accessible in free software. This advance matters because modeling dichotomous items as continuous may produce spurious factors tied to item difficulty (McDonald & Ahlawat, 1974). Accordingly, the present study incorporates more recent procedures in the analysis of the three models investigated.

2. Method

2.1. Participants

The present study used data from an open, publicly available dataset provided by Golino and Gomes (2014), comprising 292 Brazilian high school students from a public school in Belo Horizonte, Brazil, recruited through convenience sampling. This sample constitutes a subset of the 558 participants analyzed by Gomes and Borges (2009a). The sample included 53.40% females and 46.60% males, with ages between 14 and 20 years (M = 15.71, SD = 1.15). Household income was distributed across three brackets defined by the Brazilian national minimum wage of R$ 350 in 2006 (approximately USD 164 at the December 31, 2006 exchange rate of R$ 2.1372/USD): 132 participants reported income up to 5 minimum wages (approximately USD 819), 101 reported income between 5 and 10 minimum wages (approximately USD 1,638), and 40 reported income above 10 minimum wages (above USD 1,638). Household income data were missing for 19 participants (6.5%).

2.2. Instrument

The Folding Test measures the specific cognitive ability of visualization, defined as the capacity to mentally transform two-dimensional representations into three-dimensional objects (Carroll, 1993). Although the test targets this narrow ability, Carroll's (1993) evidence shows that any intelligence test is explained by three strata representing levels of breadth of cognitive abilities. The Folding Test is thus explained by the specific ability of visualization, the broad ability of visual-spatial thinking, and the general factor of intelligence. The test is part of the BAFACALO.

The full instrument and answer key can be accessed in Gomes and Araujo

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

(2025b) and Gomes et al. (2021a). It comprises 30 items organized into six pairs of figures, each pair containing five items. Each pair presents a two-dimensional figure alongside its corresponding three-dimensional object. The two-dimensional figure contains numbered dashed fold lines, and one of its faces is marked with the letter X, which designates the front face of the three-dimensional object when assembled. The sides of the three-dimensional object are labeled with letters, and the respondent must identify which letters correspond to the numbered lines on the two-dimensional figure by mentally folding it. The test has a time limit of 15 minutes. Items are scored dichotomously (0 = incorrect, 1 = correct).

2.3. Procedure

The data were published in 2014 in the Journal of Open Psychology Data, a peer-reviewed journal dedicated to data papers in psychology (Golino & Gomes, 2014). The journal accepts data papers only for datasets deposited in a public repository under an open license, so that the data are citable and available for reuse by researchers worldwide. As Golino and Gomes (2014) reported, participants completed all 18 BAFACALO tests in pencil-and-paper format across two sessions of 100 minutes each, with nine tests administered per session. Data collection complied with the guidelines of the Ethics Committee of the Universidade Federal de Minas Gerais (protocol no. ETIC 181/06). Consent was obtained from the school, participants aged 18 or older provided informed consent, and participants under 18 provided assent, with informed consent from their parents or legal guardians. Anonymity was ensured through participant identification numbers.

2.4. Data analysis

The analysis examines the dimensionality of the Folding Test through three models tested by confirmatory factor analysis. The unidimensional model specifies a single latent variable on which all items load. The Gomes and Borges (2009a) model specifies a bifactor structure with a general factor (visualization) and three specific factors: items 26–30 on the first factor; items 1–6 and 8–12 on the second factor; and items 7, 13–15, and 17 on the third factor. Factor-item relations with loadings below .20 were not included, in order to better delimit the specific factors. The relations

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

excluded from the first factor (items 1, 7, 11, 12, and 16–25) had very low loadings, most of them below .10. The relations with items 15, 16, and 18 were excluded from the second factor, and those with items 9, 20, 25, and 27 from the third. The figure model specifies the general factor in the same way as the Gomes and Borges (2009a) model, and its specific factors as follows: items 1–5 (figure 1), 6–10 (figure 2), 11–15 (figure 3), 16–20 (figure 4), 21–25 (figure 5), and 26–30 (figure 6).

Bifactor models may require adjustment, because once the general factor is modeled, specific factors can retain weak or unstable loadings, and the solution can contain improper estimates. Adjustment follows four rules: (1) in a Heywood case, the factor loading would be fixed at 1; (2) a factor-item relation that acquired a negative loading as a result of factor orthogonalization would be removed; (3) a specific factor with non-significant variance would be removed; and (4) a factor-item relation with a loading below .20 would be removed. The rules would be applied recursively, one modification at a time, with the model refitted and rechecked after each modification until no condition remained. Only the final solution of each model, after all modifications, would be reported.

All analyses are performed in R, version 4.4.1 (R Core Team, 2024). The models are estimated with the weighted least squares mean and variance adjusted estimator (WLSMV), with items treated as ordinal and modeled through thresholds (lavaan package, version 0.6-21; Rosseel, 2012). A model would be considered acceptable when the scaled comparative fit index (CFI) was ≥ .90 (Brown, 2015) and the scaled root mean square error of approximation (RMSEA) was ≤ .08 (Brown, 2015). Reliability is estimated with Cronbach's alpha and with Raykov's (2001), Bentler's (2009), and McDonald's (1999) omega coefficients (semTools package, version 0.5-5; Jorgensen et al., 2021). Acceptable models are compared by their differences in fit, with ΔCFI ≤ .002 (Meade et al., 2008) and ΔRMSEA ≤ .015 (Chen, 2007) indicating that two models do not differ meaningfully.

Unanswered items accounted for 12.74% of all responses, and 109 participants (37.3%) left at least one item unanswered. The proportion of unanswered items increased toward the end of the test, from 1.0% for item 1 to a maximum of 29.1% for item 29. Unanswered items were coded as incorrect (0), following the treatment

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

adopted in the original analyses of these data (Gomes & Borges, 2009a) and described in the dataset publication (Golino & Gomes, 2014). Because all participants had the same time to complete each test, an unanswered item indicates that the participant probably could not provide an appropriate answer within the time limit.

3. Results and Discussion

All three models showed acceptable fit (Table 1). The comparison of CFI and RMSEA indicated that the Gomes and Borges model and the figure model outperformed the unidimensional model, but that neither of them fitted the data better than the other. Comparing the structures of these two models is therefore essential to determine which of them is more appropriate.

Table 1 - Fit of the CFA models

Model	scaled χ^2	df	scaled CFI	scaled RMSEA [90% CI]
Unidimensional	663.80	405	.994	.047 [.040, .053]
Gomes and Borges model	506.82	395	.997	.031 [.022, .039]
Figure model	518.75	397	.997	.032 [.024, .040]

Both bifactor models required adjustment, and this procedure revealed that the two models have practically the same structure. In the figure model, the specific factors for figures 2, 4, 5, and 6 were removed, as were the relations of the figure 3 factor with items 11 and 12, leaving a figure 1 factor (items 1–5) and a partial figure 3 factor (items 13–15). In the Gomes and Borges model, the first specific factor was removed, as were the relations of the second factor with items 8, 9, 11, and 12 and those of the third factor with items 7 and 17, leaving the second factor with items 1–6 and 10 and the third factor with items 13–15. Table 2 shows the structure of the specific factors in both models. One specific factor of the Gomes and Borges model coincides with the partial figure 3 factor, and the other differs from the figure 1 factor only by the inclusion of items 6 and 10. This inclusion neither improves model fit nor contributes to a clearer conceptual understanding of the structure. Consequently, the figure model is the more appropriate representation of the dimensional structure of the Folding Test. The reliability analysis corroborates this conclusion: neither model shows an advantage over the other, so the inclusion of items 6 and 10 did not improve

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

measurement reliability either. The general factor (visualization) showed nearly identical reliability in the figure model and in the Gomes and Borges model (Raykov's ω = .906 and .900; Bentler's ω = .959 and .955; McDonald's ω = .979 and .974, respectively), and the specific factors showed low reliability in both (ω between .124 and .349).

Table 2 - Standardized loadings on the specific factors retained in the final solutions of the two bifactor models

Item	Figure model, Factor 3	Gomes and Borges model, Factor 3	Figure model, Factor 1	Gomes and Borges model, Factor 2
1			.629	.654
2			.387	.392
3			.345	.383
4			.471	.488
5			.328	.351
6				.315
10				.298
13	.703	.702		
14	.669	.668		
15	.289	.287		

In the figure model, all items loaded substantially on the general factor (.598 to .995), and the specific factors for figure 1 (.328 to .629) and figure 3 (.289 to .703) retained loadings above .20 (Table 3).

Table 3 - Standardized loadings of the figure model

Item	General	Figure 1	Figure 3
1	.738	.629	
2	.831	.387	
3	.793	.345	
4	.761	.471	
5	.819	.328	
6	.798		
7	.738		
8	.837		
9	.681		
10	.795		
11	.830		
12	.861		
13	.598		.703

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

14	.672	.669
15	.822	.289
16	.968	
17	.904	
18	.992	
19	.987	
20	.719	
21	.806	
22	.975	
23	.801	
24	.907	
25	.995	
26	.953	
27	.919	
28	.967	
29	.936	
30	.975	

Note. All loadings significant at $p \leq .01$.

4. Conclusion

The open science movement has established test democratization as an important agenda, making validated psychological instruments freely accessible. Free access, however, does not on its own ensure that the evidence about an instrument's dimensional structure is well calibrated. The BAFACALO has been freely accessible since 2021, with a validity evidence base accumulated over more than a decade, yet some of its factorial evidence was produced with the estimation procedures available at the time. This study reexamined the factorial structure of the Folding Test through a confirmatory comparison of models, with specific procedures for dichotomous items.

The results indicate that the Folding Test is essentially unidimensional, since all models showed acceptable fit. The model that best represents the factorial structure of the test fitted slightly better than the unidimensional model by adding a specific factor representing figure 1 and another representing part of figure 3. However, this addition corroborates the strong dominance of the general factor of the model, which represents the visualization ability. This ability showed high reliability, whereas the specific factors showed low reliability, indicating that the variance shared by the items is predominantly explained by visualization.

The study thus complements the democratization of the Folding Test with better-calibrated evidence about its dimensional structure. These findings indicate that the

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

specific factors previously identified for the Folding Test by Gomes and Borges (2009a) are better explained by the test format than by distinct abilities, a possibility that the exploratory strategy of the original study could not examine.

These findings should be interpreted in light of the study sample, which consisted of a convenience sample from a single public school in Belo Horizonte. In addition, because this sample is a subset of the sample analyzed by Gomes and Borges (2009a), the fit of the Gomes and Borges model does not constitute an independent replication of that model, whereas the unidimensional and figure models were specified independently of the data. Replication in samples from other educational and geographic contexts is therefore warranted. The next step is to apply this approach to the remaining tests of the BAFACALO.

References

Alves, F. A., Flores, R. P., Gomes, C. M. A., & Golino, H. F. (2012). Preditores do rendimento escolar: Inteligência geral e crenças sobre ensino-aprendizagem. *Revista E-Psi, 1*, 97–117. https://www.researchgate.net/publication/235642286_Preditores_do_Rendimento_Escolar_Inteligencia_Geral_e_Crencas_sobre_Ensino-Aprendizagem

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

Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., Fernandez, T., & Mandell, D. S. (2015). Free, brief, and validated: Standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice*, *22*(1), 5–19. https://doi.org/10.1016/j.cbpra.2014.02.002

Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika, 74*(1), 137–143. https://doi.org/10.1007/s11336-008-9100-1

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Bornstein, R. F. (2017). Evidence-based psychological assessment. *Journal of Personality Assessment, 99*(4), 435–445. https://doi.org/10.1080/00223891.2016.1236343

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Lawrence Erlbaum Associates.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press. https://doi.org/10.1017/CBO9780511571312

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Conselho Federal de Psicologia. (2022). *Resolução CFP nº 31/2022: Estabelece diretrizes para a realização de Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológicos (SATEPSI) e revoga a Resolução CFP nº 09/2018.* https://atosoficiais.com.br/cfp/resolucao-do-exercicio-profissional-n-31-2022-estabelece-diretrizes-para-a-realizacao-de-avaliacao-psicologica-no-exercicio-profissional-da-psicologa-e-do-psicologo-regulamenta-o-sistema-de-avaliacao-de-testes-psicologicos-satepsi-e-revoga-a-resolucao-cfp-no-09-2018

Elosua, P., & Wechsler, S. M. (2025). Attitudes toward testing in Latin America: Current situation and challenges. *International Journal of Testing, 25*(3–4), 237–254. https://doi.org/10.1080/15305058.2025.2552757

Golino, H. F., & Gomes, C. M. A. (2014). Psychology data from the “BAFACALO project: The Brazilian Intelligence Battery based on two state-of-the-art models – Carroll’s Model and the CHC model”. *Journal of Open Psychology Data*, *2*(1), e6. https://doi.org/10.5334/jopd.af

Gomes, C. M. A. (2010a). Avaliando a avaliação escolar: Notas escolares e inteligência fluida. *Psicologia em Estudo, 15*(4), 841–849. https://doi.org/10.1590/S1413-73722010000400020

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Gomes, C. M. A. (2010b). Estrutura fatorial da Bateria de Fatores Cognitivos de Alta-Ordem (BaFaCalo). *Avaliação Psicológica, 9*(3), 449–459. http://pepsic.bvsalud.org/scielo.php?pid=S1677-04712010000300011&script=sci_arttext

Gomes, C. M. A. (2011). Validade do conjunto de testes da habilidade de memória de curto-prazo (CTMC). *Estudos de Psicologia (Natal), 16*(3), 235–242. https://doi.org/10.1590/S1413-294X2011000300005

Gomes, C. M. A. (2012). Validade de construto do conjunto de testes de inteligência cristalizada (CTIC) da bateria de fatores cognitivos de alta-ordem (BaFaCAIO). *Gerais: Revista Interinstitucional de Psicologia, 5*(2), 294–316. http://pepsic.bvsalud.org/scielo.php?pid=S1983-82202012000200009&script=sci_arttext

Gomes, C. M. A., & Araujo, J. (2025a). Ciência aberta e democratização de testes: Disponibilização gratuita e aberta do Teste de Velocidade Perceptiva 2 da Bateria de Fatores Cognitivos de Alta Ordem (BAFACALO). *Revista Contemporânea, 5*(2), e7431. https://doi.org/10.56083/RCV5N2-038

Gomes, C. M. A., & Araujo, J. (2025b). Democratizando os testes de inteligência: Acesso aberto e gratuito ao Teste de Dobraduras da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO). *European Journal of Education Studies, 12*(1), 109–136. https://doi.org/10.46827/ejes.v12i1.5771

Gomes, C. M. A., & Araujo, J. (2025c). Democratizando os testes de inteligência: Acesso aberto e gratuito ao Teste de Indução da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO). *Observatório de la Economía Latinoamericana, 23*(1), 1–28. https://doi.org/10.55905/oelv23n1-136

Gomes, C. M. A., & Araujo, J. (2025d). Teste de Velocidade Perceptiva 3 da Bateria de Fatores Cognitivos de Alta Ordem (BAFACALO): Disponibilização gratuita e aberta. *Cuadernos de Educación y Desarrollo, 17*(3), e7727. https://doi.org/10.55905/cuadv17n3-031

Gomes, C. M. A., & Borges, O. N. (2009a). Propriedades psicométricas do conjunto de testes da habilidade visuoespacial. *Psico-USF, 14*(1), 19–34. https://doi.org/10.1590/S1413-82712009000100004

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Gomes, C. M. A., & Borges, O. (2009b). Qualidades psicométricas do conjunto de testes de inteligência fluida. *Avaliação Psicológica, 8*(1), 17–32. http://pepsic.bvsalud.org/scielo.php?pid=S1677-04712009000100003&script=sci_arttext

Gomes, C. M. A., de Araújo, J., Ferreira, M. G., & Golino, H. F. (2014). The validity of the Cattell-Horn-Carroll model on the intraindividual approach. *Behavioral Development Bulletin, 19*(4), 22–30. https://doi.org/10.1037/h0101078

Gomes, C. M. A., Farias, H. B., & Jelihovschi, E. G. (2022). Approaches to learning does matter to predict academic achievement. *Revista de Psicología, 40*(2), 905–933. https://doi.org/10.18800/psico.202202.010

Gomes, C. M. A., & Golino, H. F. (2012a). O que a inteligência prediz: Diferenças individuais ou diferenças no desenvolvimento acadêmico? *Psicologia: Teoria e Prática, 14*(1), 126–139. http://pepsic.bvsalud.org/pdf/ptp/v14n1/v14n1a10.pdf

Gomes, C. M. A., & Golino, H. F. (2012b). Validade incremental da Escala de Abordagens de Aprendizagem (EABAP). *Psicologia: Reflexão e Crítica, 25*(4), 400–410. https://doi.org/10.1590/S0102-79722012000400001

Gomes, C. M. A., & Golino, H. (2015). Factor retention in the intra-individual approach: Proposition of a triangulation strategy. *Avaliação Psicológica, 14*(2), 273–279. https://doi.org/10.15689/ap.2015.1402.12

Gomes, C. M. A., Marques, E. L. L., & Golino, H. F. (2014). Validade incremental dos estilos legislativo, executivo e judiciário em relação ao rendimento escolar. *Revista E-Psi, 2*, 31–46. https://revistaepsi.com/artigo/2013-2014-ano3-volume2-artigo3/

Gomes, C. M. A., & Nascimento, D. F. (2021a). *A medida da habilidade de fluência do modelo CHC: Apresentando o Teste de Fluência Ideativa 2 da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.35726.28481/1

Gomes, C. M. A., & Nascimento, D. F. (2021b). *Acesso aberto ao Teste de Fluência Figural da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO) como medida da habilidade ampla de fluência do modelo CHC de inteligência* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.15593.62564/1

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Gomes, C. M. A., & Nascimento, D. F. (2021c). *Acesso aberto e gratuito ao Conjunto de Testes de Inteligência Fluida: Teste de Raciocínio Geral da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.30509.61921/1

Gomes, C. M. A., & Nascimento, D. F. (2021d). *Acesso aberto e gratuito ao Teste de Fluência Ideativa 1 da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.24821.09442/3

Gomes, C. M. A., & Nascimento, D. F. (2021e). *Apresentando o Teste de Flexibilidade de Fechamento da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.31920.28164

Gomes, C. M. A., & Nascimento, D. F. (2021f). *Disponibilizando de forma gratuita e aberta o Teste de Memória Associativa 1 da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.29964.03201/1

Gomes, C. M. A., & Nascimento, D. F. (2021g). *Disponibilizando de forma gratuita e aberta o Teste de Velocidade Numérica da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.24114.94407/1

Gomes, C. M. A., & Nascimento, D. F. (2021h). *Medidas de inteligência cristalizada: Disponibilizando o Teste de Compreensão Verbal 2 da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.36085.09447/1

Gomes, C. M. A., & Nascimento, D. F. (2021i). *Medindo a habilidade de rapidez cognitiva do modelo CHC: Apresentando o Teste de Velocidade Perceptiva 1 da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.28564.83848/1

Gomes, C. M. A., & Nascimento, D. F. (2021j). *Projeto de acesso aberto e gratuito à Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO): O Teste de Compreensão Verbal 1 do Conjunto de Testes de Inteligência Cristalizada* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.22663.32165/1

Gomes, C. M. A., & Nascimento, D. F. (2021k). *Projeto de acesso aberto e gratuito aos testes do LAICO: Teste de Raciocínio Lógico da Bateria de Fatores*

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

Cognitivos de Alta-Ordem (BAFACALO) [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.25476.45445/1

Gomes, C. M. A., & Nascimento, D. F. (2021l). *Projeto de acesso aos testes de inteligência da BAFACALO: Teste de Compreensão Verbal 3* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.10499.84001/2

Gomes, C. M. A., & Nascimento, D. F. (2021m). *Projeto de acesso da BAFACALO: Teste de Memória Associativa 2* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.23253.14565/1

Gomes, C. M. A., & Nascimento, D. F. (2021n). *Teste de Memória Visual da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.33319.47529

Gomes, C. M. A., Nascimento, D. F., & Araujo, J. (2021a). *Acesso aberto ao Teste de Dobraduras (VZ) da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.21853.95201/2

Gomes, C. M. A., Nascimento, D. F., & Araujo, J. (2021b). *Medindo a inteligência fluida: O Teste de Indução da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO)* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.17087.84641/3

Gomes, C. M. A., Nascimento, D. F., & Araujo, J. (2021c). *Projeto de testes gratuitos e abertos do LAICO: Teste de Velocidade Perceptiva 3 da BAFACALO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.36278.42563/2

Gomes, C. M. A., Nascimento, D. F., & Araujo, J. (2021d). *Teste de Velocidade Perceptiva 2 da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO): Disponibilização aberta e gratuita aos testes de medida de rapidez cognitiva do LAICO* [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.29567.53928/1

Hernández, A., Hidalgo, M. D., Muñiz, J., & Fernández-Hermida, J. R. (2021). Diferencias de opinión sobre el uso de los test: Su relación con la edad, el género y el sector laboral. *Papeles del Psicólogo, 42*(2), 94–104. https://doi.org/10.23923/pap.psicol.2960

CC BY

REMUNOM
ISSN 2178-6925

Received: 16/09/2026 - **Accepted:** 04/10/2026
Vol: 13.21
DOI: 10.66104/00fx1964
Pages: 1-17

International Test Commission. (2017). *ITC guidelines for translating and adapting tests* (2nd ed.). https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling* (Version 0.5-5) [R package]. https://CRAN.R-project.org/package=semTools

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates.

McDonald, R. P., & Ahlawat, K. S. (1974). Difficulty factors in binary data. *British Journal of Mathematical and Statistical Psychology, 27*(1), 82–99.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568–592. https://doi.org/10.1037/0021-9010.93.3.568

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

R Core Team. (2024). *R: A language and environment for statistical computing* (Version 4.4.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology, 54*(2), 315–323. https://doi.org/10.1348/000711001159582

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Toro, S., González, I., Collao, F., Alfaro, J., Revello, V., Piñones, B., & Juliá, M. T. (2025). Significados de estudiantes de psicología sobre su formación en ética profesional. *Revista de Psicología*, *43*(1), 108–150. https://doi.org/10.18800/psico.202501.005

CC BY