

DATA PAPER

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"

Hudson F. Golino¹ and Cristiano Mauro Assis Gomes²

¹ Assistant Researcher at the Laboratory for Cognitive Architecture Mapping, Universidade Federal de Minas Gerais (UFMG), Brazil

² Adjunct Professor at the Graduate School of Psychology and Graduate School of Neuroscience, Universidade Federal de Minas Gerais (UFMG), Brazil

The BAFACALO's dataset contains the answers from 292 Brazilian high-school students from a public school to 18 cognitive tests developed to assess different hierarchical levels of intelligence structure. Most of the participants were girls (53.40%), with ages ranging from 14 to 20 years old (Mean = 15.71, Standard Deviation = 1.15) and the average monthly household income varying from R\$ 1,751 to R\$ 3,500 (Reais). The BAFACALO's tests were constructed by Gomes [1,2] based on the Educational Testing Service's Kit of Factor-Referenced Cognitive Tests [3].

Keywords: intelligence; CHC; assessment; cognition; ability; Carroll

Funding Statement: The research was funded by Fundação de Apoio à Pesquisa do Estado de Minas Gerais (FAPEMIG).

(1) Overview

Context

The Battery of Higher-Order Cognitive Factors, or the BAFACALO's project, was developed by Gomes [1,2], based on the Educational Testing Service's Kit of Factor-Referenced Cognitive Tests [3]. In the theoretical domain, the BAFACALO battery was developed to assess the general intelligence factor (g), plus six broad abilities presented in Carroll's three stratum theory [4] and the CHC model [5] in high school students. These broad abilities are: fluid intelligence (Gf) [6], crystallized intelligence (Gc) [7], short-term memory (Gsm) [8], broad visual perception (Gv) [9], fluency (Gr) and broad cognitive speediness (Gs).

Collection date(s)

2006

Background

The BAFACALO battery lies within the context of expansion and modification of intelligence tests started in the 1990's. Alfonso, Flanagan and Radwan [10] argued that until 1998 the majority of the intelligence batteries available assessed only two or three second level abilities of the CHC model. However, after 1998 the intelligence batteries were modified in order to assess four or five abilities. In Brazil, only a few intelligence batteries available enabled the assessment of the broad abilities of the CHC

model [11]. Some examples are the Woodcock-Johnson III, translated and adapted to Brazilian Portuguese by Wechsler and Schelini [12], and Schelini and Wechsler's [13] multidimensional intelligence battery for children. The BAFACALO was the first Brazilian intelligence battery developed to assess Carroll's model of intelligence and the CHC model. Around 2003, Gomes [1] translated to Brazilian Portuguese 45 tests of the Factor-Referenced Cognitive Tests Kit from Educational Testing Service. After that, Gomes developed the BAFACALO, which was inspired by his previous work. The BAFACALO was created to be applied to high school students. Its structure was investigated by Gomes [2] using EFA and SEM. Each factor retention criterion used resulted in a different factor structure. The parallel analysis identified three factors, while the Kaiser criterion suggested a four factor structure. Both, the scree plot and the maximum likelihood approach, suggested a six factor structure. The second-order general factor was identified in every solution by a second-order EFA. Each structure pointed by the EFA result was tested via SEM. The three broad factors' model with a second order general factor presented the worst fit ($\chi^2/df = 3.02$, CFI = .87, RMSEA = .08). The four broad factors' model with a second order general factor presented a χ^2/df of 2.45, a CFI of .91 and a RMSEA of .07, while the six broad factors' model with a second order g presented the best fit ($\chi^2/df = 1.39$, CFI = .98, RMSEA = .04).

(2) Methods

Sample

The sample provided in the current data paper was the original sample used in the publications cited in the background section. Despite the fact that the study included a broad variety of SES levels, which reflected the Brazilian population of high school students, the sample was not intended to be representative of this population. On the contrary, it was a convenience sample composed of 292 Brazilian high-school students from one public school (53.40% of girls, 46.60% of boys) located in Belo Horizonte, Minas Gerais, Brazil. The female majority in the sample girls, reflected a demographic characteristic of the Brazilian population. Their ages ranged from 14 to 20 years old (Mean = 15.71, Standard Deviation = 1.15). Most of the participants had monthly household incomes varying from R\$ 1,751 to R\$ 3,500 Reais. In order to recruit participants for this study, the school principal sent an invitation letter to all the school students, with the research purpose, the name and contact of the research team, as well as the dates of data collection. The chief of the school's Psychology Department visited every class reinforcing the Principal's invitation to participate in the study, and answered every question raised by the students. Those interested in being part of the study were contacted by the researchers and signed a consent form, and confirmed to be at school in the scheduled testing days. From the 320 students enrolled in the school, 91.25% accepted being part of the study, and answered the tests. It was not possible to obtain information about sampling bias in the 8.75% of students that did not agree to participate of the study. The school only dispersed data to the researchers about the students that agreed to participate in the research and signed the consent form. It is worth mentioning that the Ethics Committees in Brazil does not allow incentives in researches involving the human being, so no incentive was given to the students.

Materials

The Battery of Higher-Order Cognitive Factors (BAFACALO) was developed by Gomes [1,2].

BAFACALO broad factor's subtests

1) Fluid intelligence subtests

- Numeric Reasoning (N): The test is composed of 30 subtraction items and 30 multiplication items, with numbers containing one or two digits. The test has a time limit of two minutes.
- Inductive Reasoning (I): Composed of 15 items and a time limit of 14 minutes for its completion. Each item consists of five groups of four letters. Among the five groups there are four groups that have the same organization rule. The respondent must identify the group that has a different rule and mark it with an (x). Data code of the items: Ii01 to Ii15. The items Ii10, Ii12 and Ii13 were not included in the final models published. We discovered, after collecting the data, that these items had two correct options instead of one, so we withdraw the items from the analysis. Example presented in the instruction page:

ABCD	EFGH	IJLM	NOPQ	RSUV
------	------	------	------	------

- Logical Reasoning (RL): Composed of 30 items and a time limit of 24 minutes for its completion. Each item consists of two abstract premises and a logical conclusion based on them. The participant has to indicate whether the conclusion is logically valid or not. The participants are instructed to rely on the logical relation between the premises, not on the content itself. Data code of the items: RLi01 to RLi30. Example presented in the instruction page:

<i>All trees are fish. All fish are horses. Thus, all the trees are horses.</i>	Logically valid	Logically invalid
---	-----------------	-------------------

- General Reasoning (RG). Composed of 15 items and the completion time is limited to 18 minutes. Each item consists of a logical-mathematical problem. The respondent must interpret the item statement, solve the problem and choose one of the five possible answers. Data code of the items: RGi01 to RGi15.

2) Crystalized intelligence subtests

- Verbal Comprehension Test 1 (V1). Composed of 24 items, the time limit is set to a maximum of six minutes. Each item consists of a reference word and five multiple-choice options. Each option has one word, and the goal is to identify the word which the meaning is the closest to the reference word. Data code of the items: V1i01 to V1i24.
- Verbal Comprehension Test 2 (V2). This test has the same structure as V1, but with 18 items. Data code of the items: V2i01 to V2i18.
- Verbal Comprehension Test 3 (V3). This test has the same structure as V1, but with 18 items. Data code of the items: V3i01 to V3i18.

3) Short-term memory subtests

- Visual Memory Task (MV): The participants are requested to memorize a number of maps presented in the first sheet of the test in 3 minutes. After the memorization period, the participants have to answer to 12 items in 4 minutes. Each item is a different map, and the task is to recall if the map was presented previously in the memorization sheet. Data code of the items: MVi01 to MVi12.
- Associative Memory Task 1 (MA1): The task is to memorize 15 pairs, each pair is composed by a word and a two-digit number. The time limit is 3 minutes. After the memorization period, the participants receive a page with an unordered list of words. The task is to put the corresponding number in front of each one of the 15 words. The participants have two minutes to answer the test items. Data code of the items: MA1i01 to MA1i15.
- Associative Memory Task 2 (MA2): The task is to memorize a list with 15 names and surnames in 3 minutes. After the memorization period, the participants are given a page with an unordered list of surnames. The task is to put the corresponding name in front of

each surname. The participants have two minutes to answer the test items. Data code of the items: MA2i01 to MA2i15.

4) Broad visual perception subtests

- Visualization Test (VZ): The test consists of 30 items. Each item has a two-dimensional figure with numbered edges, and one of the sides is marked with the letter X. A three-dimensional figure created from the two-dimensional is also presented, but instead of numbers, the edges are associated with different letters. The task is to associate each edge number of the two-dimensional figure with a letter of the three-dimensional figure. Data codes of the items are: VZi01 to VZi30. An example of item is presented in **Figure 1**.
- Closure Flexibility Test (CF): The test consists of 32 items. Each item has a figure formed by four lines. A 5x5 point grid is given in the right side of the figure, and the participant is requested to recreate the figure inside the grid. The test has a time limit of 12 minutes. Code of the items in the dataset: CFi01 to CFi32. An example of item is presented in **Figure 2**.

5) Fluency subtests

- Figural Fluency test (FF): The test consists of 20 items and a time limit of one minute and a half. Each item consists of a blank T-shirt. The participant should draw details in the T-shirt, so the design on each shirt is unique. The test score consists of the number of t-shirts drawn. Test code into the dataset: FF.
- Ideational Fluency Test 1 (FI1): The participant is requested to write the greatest number of ideas related to a predetermined topic. A topic example is "A train trip." The score consists of the number of ideas written related to the topic. Test code into the dataset: FI1.
- Ideational Fluency Test 2 (FI2): The participant is requested to write the greatest number of objects related to a predetermined category. An example of category is "Red Objects." The score consists of the number of objects written related to the topic. Test code into the dataset: FI2.

6) Broad cognitive speediness subtests

- Perceptive Speed 1 (P1): Ten columns with 410 words each are presented to the participant. The task consists in marking the five words that have the letter "A" in each column. The test has 50 words with the letter "A", and a time limit of two minutes. Test code into the dataset: P1.
- Perceptive Speed 2 (P2): The test consists in 48 pairs of numbers with at least three digits. The task consists in marking all the pairs in which the numbers are different. The test has a time limit of two minutes. Test code into the dataset: P2.
- Perceptive Speed 3 (P3): The test consists in 48 items and a time limit of two and a half minutes. Each item has a target figure and five answer options. Each option has a figure, and the participant is requested to identify which one is exactly the same as the target figure. Test code into the dataset: P3.

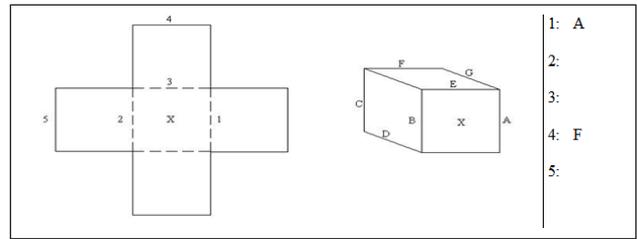


Fig. 1: Example item of Visualization Test (VZ).

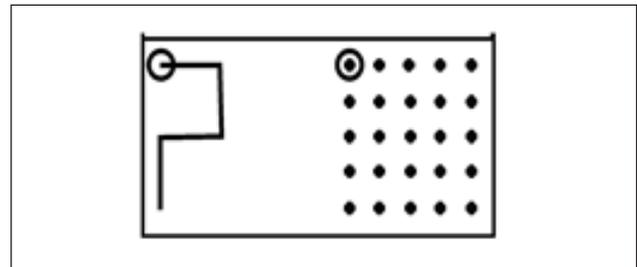


Fig. 2: Example item of Closure Flexibility Test (CF).

	Scholarly Father	Scholarly Mother	Household Income
0	4.8%	5.1%	
1	28.1%	27.5%	48.4%
2	37.7%	41.8%	37.0%
3	22.6%	25.6%	14.7%
Missing	6.9%	5.5%	6.5%

Table 1: Distribution of parental education and household income.

Procedures

The participants answered to 18 cognitive tests in two different occasions. In each occasion they had a limited time of 100 minutes to answer nine tests. The participants did not receive any incentives to answer the tests, because the Brazilian Ethical Committees do not allow this. Data related to parents schooling, household income, class number, sex and previous school type (public or private) were also collected using a personal information questionnaire previously collected by the school. The school collected this kind of data annually, as part of the school routine. Parental education was recorded as the highest educational level achieved, and is coded as follows: 0 – Pre-primary level; 1 – Primary level; 2 - Secondary level; 3 – Undergraduate level or higher levels. Household income was recorded in three categories: 1 – From 0 to 5 minimum wages; 2 – From 5 to 10 minimum wages; 3 – More than 10 minimum wages. The distribution of parental education and household income’s categories are in **Table 1**.

At the end of the year we had access to the student’s annual grades in the following subjects: Portuguese, English, mathematics, biology, physics, chemistry, geography and history. Missing data are present due to several different reasons. Some participants did not answer the personal information questionnaire provided by the school; some did not answer some of the specific tests,

or did not have time to complete a particular test. Some students did not give permission to access their annual grades, or their grades were not available.

Quality control

Participants with more than 60% of the items missing were excluded from the dataset. The final dataset have only 1% of item missingness greater than 50%. Table S1 shows the type of missing data, and how was it handled in the original publications of the test. The missing data from the first seven variables in the dataset (educational attainment of_father, educational attainment of_mother, household_income, class_number, previous_school_type, sex and age) was due to lack of information provided by the school. As pointed before, these variables were previously collected by the school, as part of the school routine. The missing data in the variables ranging from 8 to 226 in the dataset (from VZi01 to li15) was due to the participant not completing the test, i.e. he/she left some items without an answer. So, as the participants have had the same time to answer all the items, test by test, in the original publications these missing items were treated as error (i.e. received a score of 0). This strategy can be justified by the argument that in a test involving a time limit, participants who did not answer an item probably could not give an appropriate answer. On the other hand, the missing data from the variables ranging from 227 to 233 (N, P1, P2, P3, F1, F2 and FF) was due to the participant not answering to the each test in particular. So, those with missing in the variables from 227 to 233 were dropped from the analysis. Finally, the dataset codebook can also be found in Table S1 (see supporting information).

Ethical issues

Data were collected in accordance with the guidelines of the ethics committee from the *Universidade Federal de Minas Gerais*. The students interested in participating in the project signed an informed consent and waited for the researchers call. Data were anonymized by using participant identification number. Only the research staff had access to the participants' personal data.

(3) Dataset description

Object name

BAFACALO_DATASET.RData

Data type

The dataset is a 292x247 matrix containing the project's primary data. The dataset R object name is BAFACALO_DATASET.

Format names and versions

The data (an object named BAFACALO) are available as a Rdata file, created with R version 3.0.1.

Dataset creators

Hudson Golino assisted in collecting, organizing, munging and cleaning the data. Cristiano Mauro Assis Gomes designed the project, collected and analysed the data.

Language

English

License

CC-BY

Repository location

DOI: <http://dx.doi.org/10.7910/DVN/23150>

Publication date

01 November 2013

(4) Reuse potential

The data from the BAFACALO battery are useful for research of the nature of individual differences in intelligence. As we have provided the raw data, the structure of the test can be compared with different models. Furthermore, the data can be used to perform a meta-analytic structural equation modelling with other intelligence's dataset. Also, new analysis can be made investigating the role of different broad abilities in school achievement. Researchers are encouraged to use the data provided in this paper for educational purpose.

Supporting information

Table S1 (XLSX)

References

1. **Gomes, C.M.A.** (2005). *Uma análise dos fatores cognitivos mensurados pelo Exame Nacional do Ensino Médio (ENEM)*. Unpublished doctoral dissertation, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.
2. **Gomes, C.M.A.** (2010). Estrutural fatorial da Bateria de Fatores Cognitivos de Alta-Ordem (BaFaCALO). *Avaliação Psicológica*, 9(3), 449-459.
3. **Ekstrom, R. B., French, J. W., Harman, H. H. & Dirmen, D.** (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton: Educational Testing Service.
4. **Carroll, J. B.** (1993). *Human cognitive abilities: a survey of factor analytic studies*. New York: Cambridge University Press. DOI: <http://dx.doi.org/10.1017/CBO9780511571312>
5. **McGrew, K.S., Keith, T.Z., Flanagan, D.P., & Vanderwood, M.** (1997). Beyond g: the impact of Gf-Gc specific cognitive ability research on the future use and interpretation of intelligence tests in the schools. *School Psychology Review*, 26, 189-201.
6. **Gomes, C. M. A., & Borges, O. N.** (2009a). Qualidades psicométricas do Conjunto de Testes de Inteligência Fluida. *Avaliação Psicológica*, 8(1), 17-32.
7. **Gomes, C. M. A.** (2012). Validade de Construto do Conjunto de Testes de Inteligência Cristalizada (CTIC) da Bateria de Fatores Cognitivos de Alta-Ordem (BaFaCALO). *Gerias: Revista Interinstitucional de Psicologia*, 5, 294-316.
8. **Gomes, C. M. A.** (2011). Validade do Conjunto de Testes da Habilidade de Memória de Curto-Termo (CTMC). *Estudos de Psicologia (Natal)*, 16, 235-242.

9. **Gomes, C. M. A., & Borges, O. N.** (2009). Propriedades Psicométricas do Conjunto de Testes da Habilidade Visuo-Espacial. *Psico-USF*, 14(1), 19-34. DOI: <http://dx.doi.org/10.1590/S1413-82712009000100004>
10. **Alfonso, V. C., Flanagan, D. P., & Radwan, S.** (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In Dawn P. Flanagan & Patti L. Harrison (Eds.), *Contemporary Intellectual Assessment: theories, tests, and issues* (pp. 185-202). New York: Guilford Press.
11. **Primi, R.** (2003). Inteligência: avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica*, 2(1), 67-77.
12. **Wechsler, S. M., & Schelini, P. W.** (2006). Bateria de habilidades cognitivas Woodcock-Johnson III: validade de construto. *Psicologia: Teoria e Pesquisa*, 22, 287-295. DOI: <http://dx.doi.org/10.1590/S0102-37722006000300005>
13. **Schelini, P. W., & Wechsler, S. M.** (2005). Bateria multidimensional de inteligência infantil: desenvolvimento de instrumento. *Psico-USF*, 10(2), 129-139. DOI: <http://dx.doi.org/10.1590/S1413-82712005000200004>

How to cite this article: Golino, H F and 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, DOI: <http://dx.doi.org/10.5334/jopd.af>

Published: 1 August 2014

Copyright: © 2014 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License (CC-BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/3.0/>.