Avaliação Psicológica
ISSN 2175-3431 (versão on-line)

Academic Self-Report Questionnaires Measure Academic Metacognitive Knowledge: Testing this Claim

Cristiano Mauro Assis Gomes[1]
Federal University of Minas Gerais, Belo Horizonte-MG, Brazil
Enio Galinkin Jelihovschi
Santa Cruz State University, Ilhéus-BA, Brazil

ABSTRACT

Gomes and Golino (2014) made an original statement, claiming that, in essence, academic self-reported questionnaires measure the academic metacognitive knowledge. However, their evidence derived from similar observable variables based on self-report questionnaires and a small sample. In this paper we study the Gomes and Golino's assertion, using a broad sample and different types of observable variables. A favorable and two refuting models were tested, as well the role of academic metacognitive knowledge in predicting academic achievement. Only the favorable model had an acceptable data fit. Furthermore, the academic metacognitive knowledge increased incremental prediction to academic achievement. These results bring more robust evidence about the Gomes and Golino's claim.
Keywords: Students' approaches to learning; academic self-reference domains; fluid intelligence; National Exam of Upper Secondary Education (*Exame Nacional do Ensino Médio* [ENEM]); validity.

RESUMO – Os Questionários de Autorrelato Acadêmico medem, em essência, Conhecimento Metacognitivo Acadêmico: Examinando este Postulado

Gomes e Golino elaboraram um postulado original, afirmando que os questionários de autorrelato acadêmico medem, em essência, o conhecimento metacognitivo acadêmico. No entanto, suas evidências derivam de variáveis observáveis similares e uma amostra homogênea. Nosso estudo investiga o postulado de Gomes e Golino, usando uma amostra ampla, assim como tipos variados de variáveis observáveis. Um modelo favorável e dois refutadores foram testados, assim como o papel do conhecimento metacognitivo acadêmico como preditor do desempenho acadêmico. Somente o modelo favorável apresentou ajuste aceitável. Ademais, o conhecimento metacognitivo acadêmico aumentou a predição do desempenho acadêmico. Esses resultados trazem evidências mais robustas ao postulado de Gomes e Golino.
Palavras-Chave: abordagens de aprendizagem; domínios de autorreferência acadêmica, inteligência fluida; Exame Nacional do Ensino Médio (ENEM); validade.

RESUMEN – Los Cuestionarios de Autoinforme Académico miden, en esencia, el Conocimiento Metacognitivo Académico: Examinando este Postulado

Gomes y Golino elaboraron un postulado original, afirmando que los cuestionarios de autoinforme académicos miden, en esencia, el conocimiento metacognitivo académico. Sin embargo, su evidencia deriva de variables observables similares basadas en cuestionarios de autoinforme y una muestra homogénea. Este artículo investiga la afirmación de Gomes y Golino, utilizando una muestra amplia y tipos de variables observables. Se evaluó un modelo favorable y dos modelos refutadores, así como el papel del conocimiento metacognitivo académico como predictor del rendimiento académico. Únicamente el modelo favorable presentó un ajuste de datos aceptable. Además, el conocimiento metacognitivo académico aumentó la predicción incremental del rendimiento académico. Estos resultados aportan evidencias más robustas sobre la afirmación de Gomes y Golino.
Palabras Clave: Enfoques de aprendizaje; dominios de autorreferencia académica; inteligencia fluida; Examen Nacional de la Enseñanza Secundaria *(Exame Nacional do Ensino Médio* [ENEM]); validez.

The article by Gomes and Golino (2014) presented a novel argument that the constructs measured by self-report tests in the school/academic domain are all parts of a general component, called academic metacognitive knowledge (AMcK). At the operational level, this argument implies that self-report test measures of school/academic constructs are all loaded by a general factor, academic metacognitive knowledge. Gomes and Golino (2014) presented four principles to support their argument:

1. All of an individual's knowledge about their own internal processes comes from metacognition.

2. The theory of experiential structuralism states that metacognition is composed of two broad components: (1) a memory about the internal process and the self; (2) a self-regulatory component (Demetriou, 1998; Demetriou et al., 2010). The memory component of the internal process and self is all the knowledge that the individuals have about themselves and their internal

[1] Endereço para correspondência: Departamento de Psicologia. Universidade Federal de Minas Gerais. Avenida Antônio Carlos, 6627, Sala 4010, Pampulha, 31270-901, Belo Horizonte, MG. Tel.: 55 (31) 3409-5027. *E-mail:* cristianomaurogomes@gmail.com

processes. It is organized in a structure called the long-term hypercognitive system. In turn, the self-regulatory component is the cognition that regulates the individual's own cognitive processes at times when they perform any activity. It is organized in a structure called the working cognition system.

3. Constructs measured by self-report tests in the academic domain, such as student learning approaches and motivation to learn, strongly require people's knowledge of their own internal processes. This implies that they must be components of the long-term hypercognitive system.

4. Self-report tests on constructs in the school/academic domain require internal knowledge of the individual specific to this domain. Therefore, their constructs must be components of academic metacognitive knowledge, which would be part of the long-term hypercognitive system.

In conceptual terms, Gomes and Golino (2014) defined academic metacognitive knowledge as the "general academic component of long-term hypercognitive system that connects different perceptions and judgments of people about their academic abilities." (Gomes & Golino, 2014, p. 435). In operational terms, the academic metacognitive knowledge construct is precisely the general factor that explains the common variance found in self-report-based test measures of school/academic constructs.

This proposition is unique in the literature; we found no other similar proposal. It is important to point this out because there is clear recognition in the metacognition literature that metacognition is activated in any cognitive activity and therefore influences all cognitive processes. This has been clearly established by the field since its beginnings; after all, the most basic understanding of metacognition is that it is the cognition of cognition (Flavell, 1987). Therefore, the assertion that people's responses to tests involve metacognitive components is nothing new and would make no contribution to the literature. Gomes and Golino's (2014) argument is not about this. They propose that the constructs measured by self-report tests in the school/academic domain are explained in their common variance (communality) by academic metacognitive knowledge. Furthermore, this proposal is supported by conceptual and operational principles. For this reason, throughout our article we will take this as a basis and will focus on the work of Gomes and Golino (2014).

To test their argument, the study of Gomes and Golino (2014) included three constructs measured by self-report tests in the school/academic domain as observable variables. They are deep and surface approaches to learning and motivation to learn. They also included the constructs of monitoring in reading and judgment in arithmetic expressions, which come from performance-based tests and are components of the working cognition system. Gomes and Golino (2014) tested two models, one favorable and one unfavorable to their proposal. The results showed that the refuting model had an inadequate data fit ($\chi^2[8]=79.94$, CFI=.93, RMSEA=.12). The favorable model had a good data fit ($\chi^2[7]=16.52$, CFI=.99, RMSEA=.04). Since the refuting model was rejected and the favorable model showed a good data fit, Gomes and Golino (2014) evaluated the predictive validity of academic metacognitive knowledge. They found that this construct explained 6.30% of the variance of a general academic achievement, which was a latent variable that explained the variance in the annual performance of a school's students in the subjects of Portuguese, Mathematics, Geography and History. This amount in the explanation of variance was additional evidence in favor of Gomes and Golino's (2014) proposal, as the meta-analytic structural equation model of Ohtani and Hisasaka's (2018) meta-analysis showed that metacognition measured by self-report-based tests explains only 3.24% of the variance in academic performance (95% CI, 1.69% to 4.84%), an amount consistent with that found in Gomes and Golino (2014). In the meta-analytic structural equation model by Ohtani and Hisasaka (2018), metacognition and intelligence were the predictors of academic performance.

The evidences of Gomes and Golino (2014) are promising, but they arose from a narrow sample of students and one school. Furthermore, Gomes and Golino (2014) used three very similar constructs, that is, deep and surface students' approaches to learning as well as motivation for learning. The theory of learning approaches states that approaches are a combination of student strategies and motivations (Contreras et al., 2017); however, one could argue that the latent variable of Gomes and Golino (2014) was not truly the academic metacognitive knowledge but a general factor of students' approaches to learning, since motivation for learning could be a component of learning approach.

Our study aims to evaluate Gomes and Golino's claim, using a broad sample with students of many schools from two cities in the state of Minas Gerais (MG), Brazil. It also includes as observable variables six constructs measured by self-report tests in the school/academic domain. Of these, four are concepts related to the self, i.e. academic self-concept, the value attributed to academic abilities, academic self-efficacy and academic self-esteem; the remaining two are the deep approach and the surface approach to learning. Those constructs were included in order to avoid criticism of Gomes and Golino's (2014) study, since the constructs related to the self are clearly distinguished from the deep approach and the superficial approach in their conceptual definition. The former come from theories about the self (Epstein, 1973), while the latter come from the theory of learning approaches (Contreras et al., 2017). We also included as observable variables three constructs measured by performance-based tests, which are inductive reasoning, general reasoning and logical reasoning. In order to test Gomes and Golino's (2014) proposition, it is always mandatory to include constructs with measures from performance-based tests, because if academic metacognitive knowledge significantly explains the variance of the measures from these tests, then Gomes and Golino's proposition is summarily refuted.

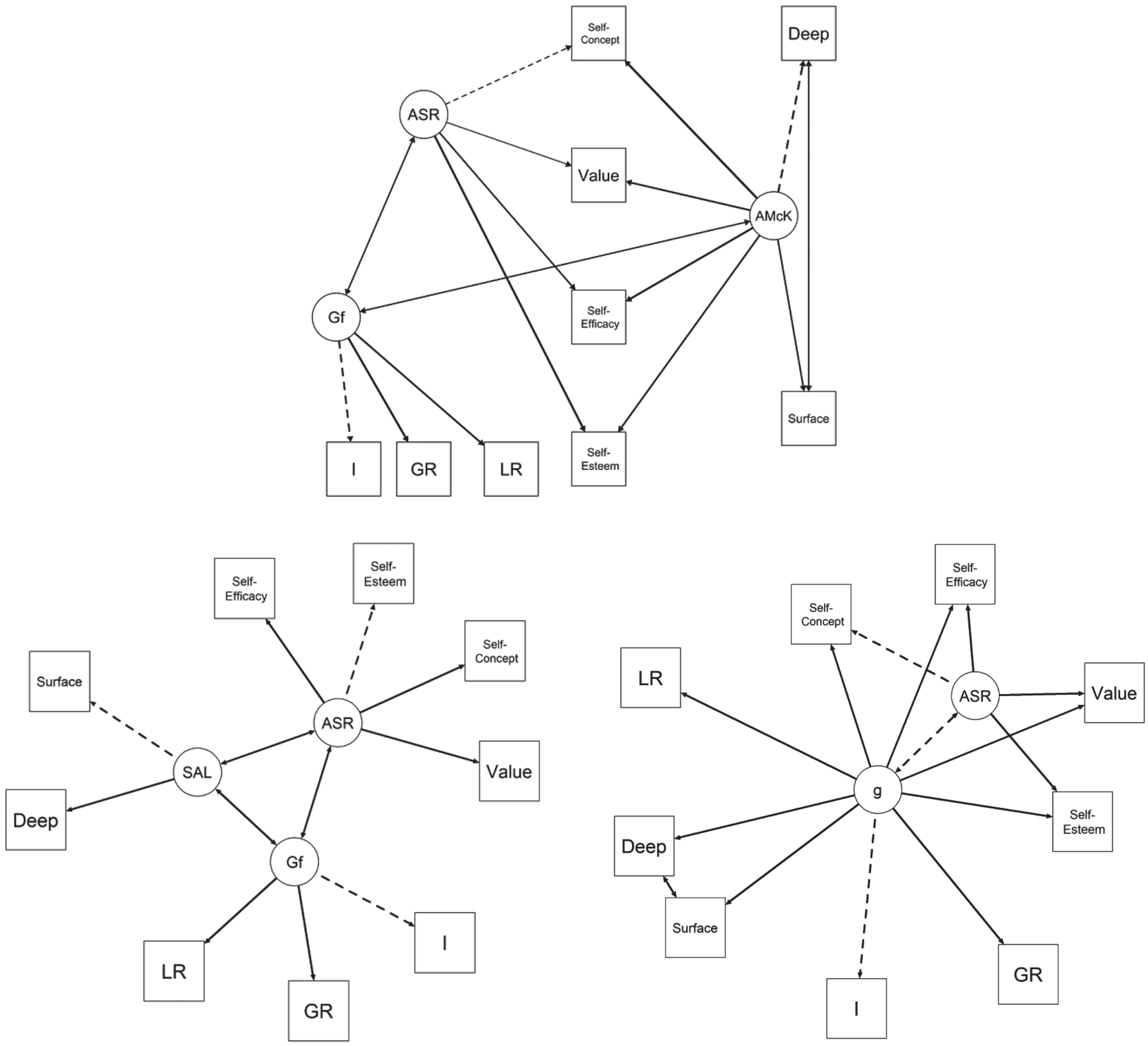
We tested a model favorable to Gomes and Golino's (2014) proposition and two models that refute it. The favorable model (top of Figure 1) assumes that academic self-esteem, the value attributed to academic abilities, academic self-efficacy, academic self-esteem, deep approach, and surface approach are all explained in their variance by the latent variable academic metacognitive knowledge (AMcK). The model also assumes that academic self-esteem, the value attributed to academic abilities, academic self-efficacy and academic value are explained in their variance by the latent variable academic self-reference, since they are all constructs about the self (Epstein, 1973). Taking as reference the evidence from Carroll's (1993) solid analysis, the model assumes that general reasoning, inductive reasoning and logical reasoning are explained in their variance by the latent variable fluid intelligence. The model also correlates the students' deep approach and superficial approach, since the proposition of the theory of learning approaches that these approaches are opposites and should correlate negatively (Contreras et al., 2017). Finally, the correlation of academic metacognitive knowledge and academic self-reference is fixed to zero.

The first refuting model assumes that the constructs of the study's observable variables are explained in their variance by the latent variables defined by current theories. In this model, the latent variable academic self-reference explains the variance of the constructs related to the self, i.e. self-concept, self-esteem, self-efficacy and value (Epstein, 1973). The latent variable fluid intelligence explains the variance of the constructs logical reasoning, inductive reasoning and general reasoning, taking as a reference the solid evidence provided by Carroll's (1993) three strata theory. Furthermore, the deep and surface approaches are explained by the latent variable general students' approach to learning, which is in line with the theory of learning approaches (Contreras et al., 2017). From the general approach to learning, it is expected that one of the approaches to positively load while the other to negatively load. This model does not assume the existence of the academic metacognitive knowledge (left of Figure 1). Furthermore, if the current theories are sufficient to explain the relationships between the constructs selected in our study, then this model should present a better fit to the data compared to the favorable model of Gomes and Golino's (2014) proposal. Considering that these theories are well-established, this refuting model represents a serious threat to that proposal.

The second refuting model assumes that a general factor explains the variance in all constructs of the model, just as the latent variable academic self-reference explains the variance in the constructs related to the self. In addition, students' superficial approach and deep approach are correlated. This model is also a strong threat to Gomes and Golino's (2014) proposal. The general factor was strategically defined in this model because it is plausible to argue against Gomes and Golino's (2014) proposal that both the constructs measured by self-report tests and the constructs measured by performance-based tests have a relevant common variance (communality), which would be explained by this general factor. If this is true, then Gomes and Golino's (2014) proposal is summarily refuted. The correlation between the general factor and academic self-reference is set to zero in this model (right of Figure 1).

An additional model will be tested in our study if the best model has an acceptable fit. This model will include students' academic achievement as an outcome variable. The model defines that some predictors explain students' academic achievement variance, which allows the vetting of what variables are relevant to predicting this outcome. We use this model to check whether the prediction results are similar to those found in Gomes and Golino (2014).

Figure 1
The favorable model and the refuting models of the Gomes and Golino's claim



Note. AMcK=academic metacognitive knowledge; ASR=academic self-reference; Gf=fluid intelligence; SAL=students' approach to learning; I=inductive reasoning; GR=general reasoning; LR=logical reasoning; Deep=deep approach; Surface=surface approach; Value=value attributed to academic abilities; g=general factor

Method

Participants

The participants of this study consisted of high school students enrolled in three public schools and one private school of Belo Horizonte city, Brazil, as well as one federal public school and one private school of Viçosa city, Brazil (N=812). The mean of the students was 16.5 years (SD=1.25) and the majority of them were females (53.45%).

Measures

Fluid Intelligence Tests Kit (*Conjunto de Testes de Inteligência Fluida*). The Fluid Intelligence Tests Kit is composed of three tests, the Inductive Reasoning Test, the General Reasoning Test, and the Logical Reasoning Test. Each of these tests is available in full, open and free of charge in Gomes and Nascimento (2021a, 2021b) and Gomes et al. (2021).

The Inductive Reasoning Test has 12 items, the General Reasoning Test has 15 items and the Logical

Reasoning Test 30 items. The Inductive Reasoning Test requires that the respondent identifies which set of letters is different from the other sets of letters. Each item of the General Reasoning Test has a statement representing a problem that must be solved by the respondent. Each item of the Logical Reasoning Test has two premises and a logical conclusion. The respondent must answer whether the logical conclusion is correct or incorrect.

The Fluid Intelligence Tests Kit is part of the Higher-Order Cognitive Factors Kit (Bateria de Fatores Cognitivos de Alta-Ordem [BAFACALO]). The BAFACALO is pioneer in Brazil to measure the Cattell-Horn-Carroll (CHC) model of intelligence (Gomes, 2010b). This battery has 18 tests measuring general intelligence (g factor) and six broad abilities of the CHC model, and has evidence of internal validity (e.g. Gomes, 2010b; Gomes & Borges, 2009) and external validity (e.g. Gomes, 2010a).

The Learning Approaches' Scale (*Escala de Abordagens de Aprendizagem* [EABAP]). The Learning Approaches' Scale test is a self-report questionnaire of deep and surface students' approaches to learning (Gomes et al., 2011) and is available at Gomes (2022). Each of the 17 items of the questionnaire has a statement about student's behavior in the context of learning in classroom and the context of studying. The respondent needs to judge the statements evaluating how much the behavior described in each statement is present in his life by using a Likert-like scale ranging from (1) not at all, to (5) entirely present. The scale shows evidence of reliability, structure validity, predictive validity and incremental validity (Gomes, 2010c, Gomes et al., 2011).

Self-Reference Academic Cognitions Scale. The Self-Reference Academic Cognitions Scale test academic self-reference constructs: academic self-concept, academic self-efficacy, academic self-esteem, and value attributed to academic abilities. Details of the test, with sample items, are given in Costa et al. (2017).

The constructs are measured through 10 testlets, which represent 10 school abilities, i.e., text comprehension, writing, habit of studying, being attentive during the classes, understanding new contents, solving mathematical equations, performing math in his head, keeping up the tasks, making oral presentations, doing exams and evaluations. Each of these testlets has a statement for the measurement of the four self-reference constructs. The test taker must answer each statement, choosing a five-points scale. Costa et al. (2017) found evidence that the Self-Reference Academic Cognitions Scale measures academic self-efficacy, academic self-concept, academic self-esteem and value attributed to academic abilities, as the scores of these constructs show reasonable reliability (alpha ranging from .69 to .79).

Academic Knowledge Exam. The Academic Knowledge Exam is composed of three forms, each with 10 items from different editions of the Brazilian National Exam of Upper Secondary Education (Exame Nacional do Ensino Médio - ENEM). The Academic Knowledge Exam corresponds to the academic knowledge items in the Academic Knowledge and Metacognition Test Booklets, which are presented in detail in Costa (2018).

Each form of the Academic Knowledge Exam has items with very similar levels of difficulty. Each form has some common items and unique items as follows: form 1 is composed of items 1 to 10; form 2 has items 1, 2, 6, 8, 10, 11, 13, 14, and 15; and form 3 has items 1, 5, 6, 10, 13, 15, 16, 17, 18, and 19. The items were selected from the ENEM editions ranging from 2001 to 2007. Each form is composed of two very easy questions (80% to 100% of students hit the item), two easy questions (60% to 80% of hits), two median items (40% to 60% of hits), difficult items (20% to 40% of hits), and very difficult items (0% to 20% of hits).

Each of the forms of the Academic Knowledge Exam measures exclusively the general factor of students' academic performance. In the study by Gomes and Golino (2014), they used the students' annual school grades in Mathematics, Portuguese, History and Geography as observable variables to measure the latent variable of academic performance since the participants in their study came from a single school. In our study, it would not be interesting to use school grades, as schools usually have very different criteria for scoring students. An annual score of 60 points in a particular subject at one school may correspond to a score of 80 points in the same subject at another school. As the forms of the Academic Knowledge Exam are made up of well-designed items from the ENEM and have very similar levels of difficulty, they are suitable to be applied to students from different schools.

Procedures

The research was approved by the Ethics Committee of the *Universidade Federal de Minas Gerais* (UFMG), Brazil, n. 364.253. The tests were applied in the classroom by a psychologist and students of Psychology which were trained by the psychologist. The teachers were in the classroom at the moment of application of the tests. Data was collected in two applications lasting approximately 90 minutes each. Since the large number of tests and the long time needed to complete them, the participants answered only one of the three forms of the Academic Knowledge Exam in order to avoid fatigue or disinterest, considering that these three forms were designed to exclusively measure the same construct. Each form of the exam was randomly assigned to each student. The data collection procedure is presented in detail in Costa's work (2018).

Data Analysis

The data analysis followed two steps. In the first step, we tried to create model-based scores, so as not to use raw scores for the observable variables in our study. To create these scores, we first applied a measurement model to each test applied in this study. These measurement models were tested via item confirmatory factor analysis. As these models had exclusively categorical data, we used the weighted least square mean and variance (WLSMV) estimator. If the measurement model presented minimally acceptable data fit (Comparative Fit Index [CFI] ≥.90; Root Mean Square Error Approximation [RMSEA] <.10; Thakkar, 2020), we then ran a measurement model similar to the one used in the item confirmatory factor analysis, now applying full information confirmatory factor analysis. In this analysis, when the measurement model involved dichotomous items, we used the Rasch Model to generate the scores; when the measurement model involved polytomous items, we used the Partial Credit Model, as it also constrains the difficulty parameter of the items to the value 1, in the same way as the Rasch Model. A controversial feature of this model is that it does not require that the test response categories have exactly the same distance on the logit scale as the score produced by the model. This makes it difficult to easily interpret the response categories, since it is necessary to check how the categories of each item are located on the logit scale of the model score. In a model where it is assumed that the distances between the response categories are the same for all items, this inspection is not necessary. Although this makes it difficult to easily interpret the response categories of the items, it does not make the quality of the model scores inadequate.

An item confirmatory factor analysis model was applied in the items of EABAP, defining that the surface and the deep approach latent variables explained their respectively marker items. Only if this model showed an acceptable data fit (Comparative Fit Index [CFI] ≥.90; Root Mean Square Error Approximation [RMSEA] <.10; Thakkar, 2020), then a multidimensional Partial Credit Model was performed in the items of EABAP, generating the expected a-posteriori (EAP) scores for the deep and the surface approaches to learning. It was applied an item confirmatory factor analysis model in the items of the Self-Reference Academic Cognitions Scale. This model assumed that the latent variables of academic self-concept, academic self-esteem, academic self-efficacy, and value assigned to academic abilities explained their marker items. In addition, this model assumed that 10 specific latent variables explained the items of each testlet of the test. These specific latent variables did not correlate with the other factors and with themselves, excepting the correlation between specific factor one and factor two, as well specific factor six and factor seven. Making it simpler, a multidimensional Partial Credit Model was applied in the items of this test containing only the latent variables of academic self-concept, academic self-esteem, academic self-efficacy, and value assigned to academic abilities. Then, an item confirmatory factor analysis model was applied in the items of the Fluid Intelligence Tests Kit, defining that the latent variables of inductive reasoning, general reasoning, and logical reasoning explained their respective marker items. A multidimensional Rasch model containing these latent variables was applied. A one-dimensional item confirmatory factor analysis model was applied in each form of the Academic Knowledge Exam, assuming that a general academic achievement explained all the items in each test form. In addition, a unidimensional Rasch model was applied in each test form, as well on all items in all forms of the Academic Knowledge Exam. Considering that only the score from this last model with all the items from all the Exam forms would be used in step 2 of this study, only the marginal reliability of this model was calculated. Because these forms have few items in common, it is not possible to run an item confirmatory factor analysis on this model, nor to calculate the infit from the full information confirmatory factor analysis. For this reason, we have taken the results of the model adjustment and infit of the Exam forms separately as references.

The EAP scores from these analyses were evaluated through the infit of the items and the marginal reliability of the scores. The infit of the items having values between 0.5 and 2.0 and the marginal reliability equal or above .60, indicated sufficient reliability of the scores. The infit values between 0.5 and 2.0 were defined with reference to Linacre's guidelines (2002). He states that values in this range do not degrade the quality of the measurement scores. Considering that the marginal reliability is an estimate of the overall reliability of a test and that the "marginal reliability can be interpreted much like coefficient alpha but estimates reliability using information from the estimated IRT model" (Harris et al., 2020, p. 520), we used a criterion similar to that applied to Cronbach's alpha, where values equal or greater than .60 are considered minimally sufficient (Ursachi et al., 2015).

The second step involved the evaluation of the favorable model and the evaluation of the two refuting models of the Gomes and Golino's claim. These models were previously presented in the introduction to this article and are shown in Figure 1. All models in the second step were tested through structural equation models. The estimator used was the maximum likelihood estimator (ML). The scores in the first step of the analysis were used as the observable variables of the three models. These scores are shown in Figure 1 in rectangle format, representing the observable variables of the models: the scores for inductive reasoning (I), general reasoning (GR), logical reasoning (LR), deep approach (Deep), surface approach (Surface), academic self-efficacy, academic self-esteem, academic self-concept, and value assigned to academic abilities (Value). In all models there are latent variables that explain the variance of the observable variables. The latent variables are represented in Figure 1 in the shape of a circle. The model favorable to Gomes and Golino's (2014) proposition is the only one that has the latent variable academic metacognitive knowledge (AMcK). The two unfavorable models do not assume the presence of this latent variable.

When a latent variable loaded the entire set of observable variables also carried by another latent variable in the model, these two latent variables were orthogonalized to each other, i.e. their correlation was kept at zero. As a result, they both competed to explain the variance of these items. This occurred in the favorable model, where academic metacognitive knowledge (AMcK) loads on all observable variables loaded by the latent variable academic self-reference (ASR). The same occurred in the second unfavorable model, where the latent variable (g) loads on all the observable variables loaded by the latent variable academic self-reference (ASR).

The data fit of the models was assessed by the Comparative Fit Index (CFI) and the Root Mean Square Error Approximation (RMSEA). CFI values equal or above .90 and RMSEA values equal to or smaller than .10 indicated that the model should not be refuted (Thakkar, 2020). If only one model showed an acceptable fit to the data, it was automatically evaluated as the best model out of the three models tested. If more than one model showed an acceptable fit, they would be compared using the Satorra and Bentler (2001) test.

A new model would be tested, via structural equation modeling, if the best model with respect to the data showed an acceptable fit. It would have added students' academic performance as an outcome variable. If any predictor in the model had a loading in relation to the outcome variable with a p-value equal to or greater than .05, then this loading would be constrained to zero and a new model would be run.

The Rasch and Partial Credit models via full information confirmatory factor analysis were performed through the mirt R package (Chalmers, 2012), while the item confirmatory factor analyzes and the structural equation modeling were performed through the lavaan R package (Rosseel, 2012) and semTools R package (Jorgensen et al., 2020). The marginal reliability was performed through the mirt R package for the unidimensional Rasch models. This reliability can be applied only to unidimensional models in this package; for multidimensional Rasch models, the empirical reliability is applied, which is a variant of marginal reliability.

Results and Discussion

The aim of the entire first part of the analysis was solely to produce model-based scores, which would be used as the observable variables of the models favorable and unfavorable to Gomes and Golino's (2014) proposal. Table 1 summarizes the results of this analysis. All the measurement models tested via item confirmatory factor analysis showed an acceptable fit and the scores produced in the models via full information confirmatory factor analysis showed a satisfactory fit and sufficient reliability (Table 1).

Table 1
Degree of Fit of the Models, Infit and Reliability of the Measures Generated

	Goodness of data fit (item confirmatory factor analysis)	infit (full information confirmatory factor analysis)	marginal reliability (full information confirmatory factor analysis)
EABAP	χ^2[118]=616.62, CFI=.956; RMSEA=.073 (90% CI, .067 to.079)	M=.90, MIN=0.76, MAX=1.08	.78 (surface approach), .82 (deep approach)
Self-Reference Academic Cognitions Scale	χ^2[692]=3840.50, CFI=.980; RMSEA=.076 (90% CI, .074 to .078)	M=0.84, MIN=0.61, MAX=1.19	.75 (self-concept), .72 (self-esteem), .73 (self-efficacy), .74 (value)
Fluid Intelligence Tests Kit	χ^2[1536]=4130.71, CFI=.902; RMSEA=.046 (90% CI, .044 to .048)	M=0.95, MIN=0.69, MAX=1.34	.69 (inductive reasoning), .71 (logical reasoning), .78 (general reasoning)
Academic Knowledge Exam (form 1)	χ^2[35]=46.15, CFI=.904; RMSEA=.027 (90% CI, .000 to .046)	M=0.90, MIN=0.73, MAX=1.06	-
Academic Knowledge Exam (form 2)	χ^2[35]=36.07, CFI=.984; RMSEA=.018 (90% CI, .000 to .079)	M=1.02, MIN=0.72, MAX=1.35	-
Academic Knowledge Exam (form 3)	χ^2[35]=21.70, CFI=1.000; RMSEA=.000 (90% CI, .000 to .000)	M=0.93, MIN=0.59, MAX=1.26	-
Academic Knowledge Exam	-	-	.76 (academic knowledge)

Note. M=mean; MIN=minimum; MAX=maximum; χ^2=chi-square; CI=confidence interval; CFI=Comparative Fit Index; RMSEA=Root Mean Square Error Approximation

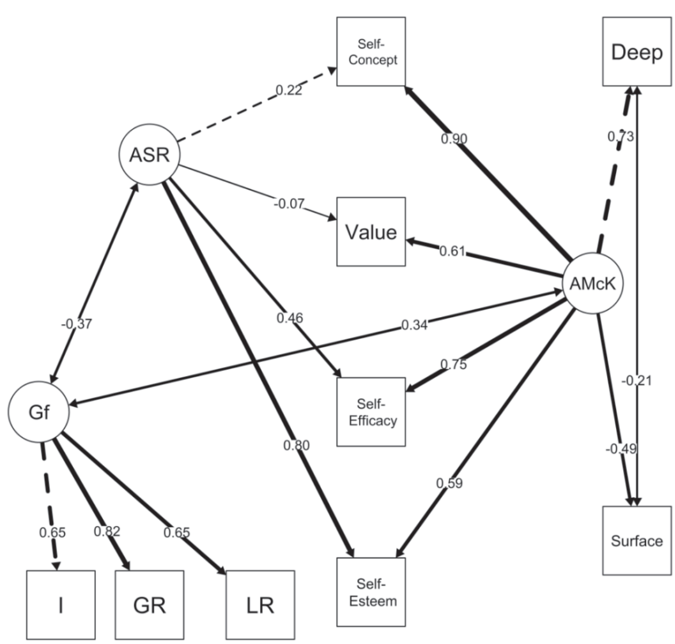
The central part of our study involves testing the favorable and unfavorable models to Gomes and Golino's (2014) proposal. The unfavorable models present assumptions capable of summarily refuting the proposal. The first refuting model is a particularly strong threat to Gomes and Golino's (2014) proposal, as it is based on well-established theories about the observable variables used in our study. We present the results of testing these models. The favorable model had a good data fit (χ^2[21]=50.12, CFI=.992; RMSEA=.042, 90% CI lower=.027, upper=.056), while the two refuting models had inacceptable data fit and must be rejected (refuting model 1: χ^2[24]=424.37, CFI=.883; RMSEA=.144, 90% CI lower=.132, upper=.156; refuting model 2: χ^2[22]=587.16, CFI=.835; RMSEA=.179, 90% CI lower=.166, upper=.191).

This result is quite surprising, as none of the unfavorable models showed a minimally acceptable fit, either in terms of CFI or RMSEA. This indicates that these models should be rejected, as they do not correctly explain the relationships between the observable variables. The favorable model not only has a minimally acceptable fit, but shows a good fit to the data. This is what the psychometric literature indicates for models with CFI ≥ .95 and RMSEA <.06 (Hu & Bentler, 1999). In short, the unfavorable models lack something that only the favorable model has, the latent variable academic metacognitive knowledge. It explains the common variance of all the observable variables from self-report tests on school/academic domain constructs. This common variance is not adequately explained by current theories; if it were, refutative model 1 should have shown a good fit to the data and should also have outperformed the fit of the favorable model. None of this happened.

Figure 2 shows the factor loadings and correlations of the favorable model. They have statistically significant values (p<.001), with the exception of the factor loading of the latent variable academic self-reference (ASR) on the observable variable value attributed to academic abilities (Value), with a p-value of .087.

All factor loadings of the latent variable academic metacognitive knowledge (AMcK) are equal to or greater than .49 (value in module). The average of these loadings, in module, is quite high (M=.68, SD=.14). This indicates that this latent variable plays an important role in explaining the variance in the scores of all these constructs measured by self-report tests. Note that this occurs both in the constructs related to the self in the academic domain and in the constructs on learning approaches.

Figure 2
Favorable Model, Loadings and Correlations



Note. AMcK=academic metacognitive knowledge; ASR=academic self-reference; Gf=fluid intelligence; I=inductive reasoning; GR=general reasoning; LR=logical reasoning

Two additional results stand out. The first shows that academic metacognitive knowledge (AMcK) has much higher factor loadings on three of the four observable variables about the self, when compared to the loadings of academic self-reference on these same observable variables. These are the value attributed to academic abilities, academic self-concept and self-efficacy. In theory, it should be the other way around, since these constructs are components of self-reference (Epstein, 1973). The second result shows that there is a positive correlation (.34) between fluid intelligence (Gf) and academic metacognitive knowledge (AMcK), as well as a negative correlation (-.37) between fluid intelligence (Gf) and academic self-reference (ASR). The meta-analytic structural equation model by Ohtani and Hisasaka (2018) shows that metacognition assessed by self-report tests and intelligence have a correlation of .20 (95% CI, .12 to .28). This result strengthens Gomes and Golino's (2014) proposal, as both results are similar. At the same time, the negative association found between fluid intelligence (Gf) and academic self-reference (ASR) may have occurred because the latter was orthogonalized to the latent variable academic metacognitive knowledge. When a model controls for the association between academic self-reference and fluid intelligence via academic metacognitive knowledge, it is possible that the real association between academic self-reference and intelligence is identified and is shown to be negative. This seems to be the case, since in unfavorable model 1, which contained the latent variable academic self-reference, the latent variable fluid intelligence and the general learning approach factor, both correlated with each other, the correlation between fluid intelligence and academic self-reference was positive (.14, *p-value*=.001). Further studies are obviously needed to verify the consistency of this result.

Another relevant result is the negative factor loading between students' superficial approach (Surface) and academic metacognitive knowledge (AMcK). This result was expected, as the theory of learning approaches states that the surface approach is the opposite of the deep approach (Contreras et al., 2017). While the former leads to poorer quality learning, the latter leads to better quality learning. For that reason, we interpret that the negative loading between academic metacognitive knowledge and the surface approach means that academic metacognitive knowledge is a type of knowledge that leads the individual to an improvement in their performance and personal fulfillment, while the surface approach leads the individual to a counter-productive practice.

The last model of this study has the same observable and latent variables, as well the relations determined by the favorable model, adding a new observable variable that is students' academic achievement. This model defines that academic metacognitive knowledge, academic self-reference, and fluid intelligence, all of them latent variables, explain students' academic performance variance. After running this model, the latent variable academic self-reference (ASR) presented a loading with a p-value greater than .05 in relation to the outcome and a new model was run with this loading constrained to zero. The new model had a good data fit ($\chi^2[28]=53.39$, CFI=.993; RMSEA=.034 (90% CI, .019 to .047).

The students' academic achievement variance was explained at 48.1%. The loading of fluid intelligence on the outcome was equal to .668 (*p-value*<.000), explaining 44.62% of its variance. In addition, the loading of academic metacognitive knowledge on the outcome was equal to .065 (*p-value*=.049), showing a small increment of the outcome variance at 0.42%. The prediction of each of these latent variables on the outcome, when added together, resulted in 45.04% of the variance being explained. The remaining 3.06% of the 48.1% comes from the correlation between fluid intelligence and academic metacognitive knowledge.

The results indicate that fluid intelligence is responsible for almost the entire explained variance of the outcome. It is important to mention that the small increase in prediction provided by academic metacognitive knowledge is consistent with the results found in the meta-analytic structural equation model by Ohtani and Hisasaka (2018). Their results show that the prediction of metacognitive tests based on self-report is low in relation to academic performance.

Although the predictive result of our study corroborates the proposal by Gomes and Golino (2014), it is important to discuss it. The students' academic achievement was measured by an academic exam which just contain items of different editions of the National Exam of Upper Secondary Education (ENEM), which is the Brazilian large-scale educational test that measures the academic knowledge of the Brazilian students in secondary education. Maybe this context explains why academic metacognitive knowledge had an increment of 0.4% in this study and in the Gomes and Golino (2014) study had an increment of 6.25% of the students' academic achievement. Gomes and Golino' used the z score of four annual school grades as observable variables, as well as, a general school achievement latent variable that explained the variance of these observable variables. The z score was calculated by school grade level.

However, there are other plausible possibilities that explain why academic metacognitive knowledge in Gomes and Golino's study had a higher incremental impact in the prediction of the outcome. Gomes and Golino had added in their model variables of working metacognition, and in this article, we added variables of reasoning and the fluid intelligence latent variable. It is possible that fluid intelligence has a more noticeable place to explain the students' academic achievement in secondary education than working metacognition. The former imposes a more restrict control in the explanation of academic metacognitive knowledge in relation to academic achievement. Other plausible explanation is that our article had a heterogeneous sample, where the students come from different schools of two Brazilian cities. Despite the size of the Gomes and Golino's (2014) sample being similar to the sample size of this article, the sample of the former comes from only one school. Another possibility, no less important, is that the difference in results is merely random fluctuation in the true result.

Conclusion

This article addressed the Gomes and Golino's claim that all academic self-report questionnaires measure, in essence, academic metacognitive knowledge, which "is a general academic component of long-term hypercognitive system that connects different perceptions and judgments of people about their academic abilities." (Gomes & Golino, 2014, p. 435).

The study by Gomes and Golino (2014) had a weakness. In the favorable model tested, academic metacognitive knowledge was a latent variable that explained the commonality of the observable variables superficial approach, deep approach, and motivation to learn. It can be argued that the latent variable academic metacognitive knowledge is simply a general factor of students' learning approach, since the theory of learning approaches states that learning approaches are a combination of strategies and student motivations. Our study selected other variables in order to solve this problem. In it, the favorable model, academic metacognitive knowledge explains the commonality of the observable variables academic self-concept, academic self-esteem, academic self-efficacy, value attributed to academic abilities, deep approach and superficial approach. In this study, it is not sustainable to speculate that the variables related to the self are components of the learning approach, so the characterization of the latent variable academic metacognitive knowledge becomes more solid in the presence of the observable variables used.

The two models that refuted Gomes and Golino's (2014) proposal were carefully selected. In the first refuting model, the latent variables included were supported by current theories about the observable variables used in the study. That is, a latent variable of general approach to learning for the observable variables of deep and surface approach, a latent variable of academic self-reference for the observable variables related to the self, and a latent variable of fluid intelligence for the observable variables of reasoning. We would expect this model to have a good fit to the data and it was surprising that this model showed fit indexes below the minimally acceptable. If this model followed current theories related to the constructs represented by the observable variables, what was missing? The answer is academic metacognitive knowledge.

In Gomes and Golino's proposal, only tests based on self-report have a common factor, which is the academic metacognitive knowledge. If this latent variable also loaded the tests based on reasoning performance, then Gomes and Golino's (2014) proposal would be summarily refuted. For this reason, the second refuting model was important to add to the study.

Nevertheless, the evidence found in this article corroborates the Gomes and Golino's claim. Only the favorable model of the claim showed good data fit, while the two refuting models had inacceptable data fit.

The Gomes and Golino's study was the first attempt to show their claim and bring initial evidence to it; this article brings more robust evidence in favor of the claim, indicating that it seems plausible and capable to explain why these different constructs from different traditions of psychology are related to each other.

Agradecimentos

Agradecimento ao CNPq, bolsa de produtividade de Cristiano M. A. Gomes. Agradecimento à Pró-Reitoria de Pós-Graduação da Universidade Federal de Minas Gerais.

Financiamento

A presente pesquisa não recebeu nenhuma fonte de financiamento sendo custeada com recursos do próprio autor.

Contribuições dos autores

Declaramos que todos os autores participaram da elaboração do manuscrito.

Disponibilidade de dados e materiais

Todos os dados e sintaxes gerados e analisados durante esta pesquisa serão tratados com total sigilo devido às exigências do Comitê de Ética em Pesquisa com Seres Humanos. Porém, o conjunto de dados e sintaxes que apoiam as conclusões deste artigo estão disponíveis mediante razoável solicitação ao autor principal do estudo.

Conflitos de interesses

Os autores declaram que não há conflitos de interesses.

References

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

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory package for the R environment. *Journal of Statistical Software, 48*(6), 1-29. https://www.doi.org/ 10.18637/jss.v048.i06

Contreras, M. S., Salgado, F. C., Hernández-Pina, F. & Hernández, F. M. (2017). Enfoques de aprendizaje y enfoques de enseñanza: Origen y evolución. *Educación y Educadores, 20*(1), 65-88. https://www.doi.org/10.5294/edu.2017.20.1.4

Costa, B. C. G. (2018). *Caminhos para predição do desempenho acadêmico: um modelo de variáveis cognitivas e socioemocionais.* [Tese de Doutoramento, Universidade de Brasília]. Repositório da Universidade de Brasília. http://www.pgpds.unb.br/images/Tese_de_Doutorado_Bianca_Cristine_Gomide_Costa.pdf

Costa, B. C. G., Gomes, C. M. A., & Fleith, D. S. (2017). Validade da Escala de Cognições Acadêmicas Autorreferentes: autoconceito, autoeficácia, autoestima e valor. *Avaliação Psicológica, 16*(1), 87-97. https://www.doi.org/10.15689/ap.2017.1601.10

Demetriou, A. (1998). Nooplasis: 10 + 1 postulates about the formation of mind. *Learning and Instruction, 8*(4), 271-287. https://doi.org/10.1016/S0959-4752(97)00022-4

Demetriou, A., Spanoudis, G., & Mouyi, A. (2010). A three-level model of the developing mind: Functional and neuronal substantiation. In M. Ferrari & L. Vuletic (Eds.), *The developmental relations between mind, brain, and education: Essays in honor of Robbie Case* (pp. 9-48). New York: Springer.

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist, 28*(5), 404-416. https://www.doi.org/10.1037/h0034679

Flavell, J. (1987). Speculations about the nature and development of metacognition. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21-29). Hillsdale, NJ: Lawrence Erlbaum.

Gomes, C. M. A. (2010a). Avaliando a avaliação escolar: notas escolares e inteligência fluida. *Psicologia em Estudo, 15*(4), 841-849. http://www.redalyc.org/articulo.oa?id=287123084020

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?script=sci_arttext&pid=S1677-04712010000300011&lng=pt.

Gomes, C. M. A. (2010c). Perfis de Estudantes e a relação entre abordagens de aprendizagem e rendimento Escolar. *Psico (PUCRS. Online), 41*(4), 503-509. http://revistaseletronicas.pucrs.br/ojs/index.php/revistapsico/article/view/6336

Gomes, C. M. A. (2022). *Projeto de Acesso Aberto: Escala de Abordagens de Aprendizagem (EABAP).* OSF Preprint. https://doi.org/10.31219/osf.io/4edy7

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

Gomes, C. M. A., & Golino, H. F. (2014). Self-reports on students' learning processes are academic metacognitive knowledge. *Psicologia: Reflexão e Crítica, 27*(3), 472-480. https://www.doi.org/10.1590/1678-7153.201427307

Gomes, C. M. A., Golino, H. F., Pinheiro, C. A. R., Miranda, G. R., & Soares, J. M. T. (2011). Validação da Escala de Abordagens de Aprendizagem (EABAP) em uma amostra brasileira. *Psicologia: Reflexão e Crítica, 24*(1), 19-27. https://www.doi.org/10.1590/S0102-79722011000100004

Gomes, C. M. A., & Nascimento, D. F. (2021a). 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*. https://doi.org/10.13140/RG.2.2.30509.61921/1

Gomes, C. M. A., & Nascimento, D. F. (2021b). Projeto de acesso aberto e gratuito aos testes do LAICO: Teste de Raciocínio Lógico da Bateria de Fatores Cognitivos de Alta-Ordem (BAFACALO). *Preprint*. https://doi.org/10.13140/RG.2.2.25476.45445/1

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

Harris, A. M., McMillan, J. T., & Carter, N. T. (2020). Test-taker reactions to ideal point measures of personality. *Journal of Business and Psychology, 36*, 513-532. https://doi.org/10.1007/s10869-020-09682-8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. https://doi.org/10.1080/10705519909540118

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2020). semTools: Useful tools for structural equation modeling. R package version 0.5-3. https://CRAN.R-project.org/package=semTools

Linacre, J. (2002) What do infit, outfit, mean-square, and standardization mean? *Archives of Rasch Measurement, 16*, 871-882. https://www.rasch.org/rmt/rmt162f.htm

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*(2), 179-212. https://doi.org/10.1007/s11409-018-9183-8

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

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. https://doi.org/10.1007/BF02296192

Thakkar, J. J. (2020). *Structural Equation Modelling Application for Research and Practice (with AMOS and R).* Springer: Singapore.

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance, 20*, 679-686. https://doi.org/10.1016/S2212-5671(15)00123-9

recebido em abril de 2021
aprovado em agosto de 2024

Sobre os autores

Cristiano Mauro Assis Gomes is Professor of Universidade Federal de Minas Gerais. Head of Laboratory for Cognitive Architecture Mapping (LAICO)/Federal University of Minas Gerais/Brazil.

Enio Galinkin Jelihovschi is Professor of Santa Cruz University.

Como citar este artigo

Gomes, C. M. A., & Jelihovschi, E. G. (2024). Academic Self-Report Questionnaires Measure, in Essence, the Academic Metacognitive Knowledge, Testing this Claim. *Avaliação Psicológica, 23*(4), 369-379. http://dx.doi.org/10.15689/ap.2024.2304.01