Measuring students' learning approaches through achievement: structural validity of SLAT-Thinking

Cristiano Mauro Assis Gomes. Universidade Federal de Minas Gerais
Josiany Soares Quadros. Universidade Federal de Minas Gerais
Jhonys de Araujo. Universidade Federal de Minas Gerais
Enio Galinkin Jelihovschi. Universidade Estadual de Santa Cruz

Abstract

The field of learning approach studies measures its constructs exclusively using self-report instruments. Recently created, SLAT-Thinking is the first test to measure performance approaches in a task that requires the respondent to identify the author's thinking in a given text. This paper presents the first evidence on the structural validity of SLAT-Thinking. A sample of 622 higher education students was randomly divided into training sample and test sample. Two models were tested in the training sample through confirmatory factor analysis of items, resulting in a final model. This model presented configural, metric, and scalar invariance when comparing the training and test sample. The results indicate that SLAT-Thinking reliably measures three levels of learning approaches: superficial, intermediate-deep, deep. The measurement of levels generates information not previously identified by the area, bringing conceptual implications for the construct.

Keywords: students' learning approach; deep approach; surface approach; structural validity; achievement test.

Resumo

Medindo Abordagens de Aprendizagem pelo Desempenho: Validade Estrutural do TAP-Pensamento. O campo de estudos em abordagens de aprendizagem mensura exclusivamente seus construtos usando instrumentos de autorrelato. Criado recentemente, o TAP-Pensamento é o primeiro teste a mensurar abordagens via o desempenho em uma tarefa que demanda ao respondente identificar o pensamento do autor em determinado texto. O presente artigo apresenta as primeiras evidências sobre a validade estrutural do TAP-Pensamento. Uma amostra de 622 estudantes do ensino superior foi dividida aleatoriamente em amostra-treino e amostra-teste. Dois modelos foram testados na amostra-treino, via análise fatorial confirmatória de itens, resultando em um modelo final. Este modelo apresentou invariância configural, métrica, e escalar, ao se comparar a amostra treino e teste. Os resultados indicam que o TAP-Pensamento mensura de forma confiável três níveis de abordagens de aprendizagem: superficial, intermediária-profunda, profunda. A mensuração de níveis gera informações anteriormente não identificadas pela área, trazendo implicações conceituais para o construto.

Palavras-chave: abordagem de aprendizagem; abordagem profunda; abordagem superficial; validade estrutural; teste de desempenho.

Resumen

Midiendo los Enfoques de Aprendizaje mediante el Desempeño: Validez Estructural del TAP-Pensamiento. El campo de estudios sobre los enfoques de aprendizaje mide sus constructos exclusivamente utilizando instrumentos de autorreporte. Recientemente creado, el TAP-Pensamiento es el primer test que mide los enfoques a través del desempeño de una tarea que demanda que la persona identifique el pensamiento del autor en un texto determinado. Este artículo presenta las primeras evidencias sobre la validez estructural del TAP-Pensamiento. Una muestra de 622 estudiantes de educación superior fue dividida aleatoriamente en muestra-entrenamiento y muestra-prueba. Dos modelos fueron testados en la muestra de entrenamiento a través del análisis factorial confirmatorio de ítems, lo que resultó en un modelo final. Este modelo presentó invarianza configural, métrica y escalar al comparar las muestras de entrenamiento y prueba. Los resultados indican que el TAP-Pensamiento mide de manera confiable tres niveles de enfoques de aprendizaje: superficial, intermedio-profundo y profundo. La medición por niveles genera información no identificada previamente por el área, trayendo implicaciones conceptuales para el constructo.

Palabras clave: enfoque de aprendizaje; enfoque profundo; enfoque superficial; validez estructural; test de desempeño.

The theory of learning approaches, in its original and relevant contributions to the field of educational psychology and psychological assessment (Gomes, 2013), brings arguments and evidence that broaden the understanding of the subject-object interaction of knowledge and its repercussion in the learning process (Gomes, 2010b). Its originality comes to light in showing that the subject's interactions with knowledge objects are guided by two fundamental approaches: one deep and one superficial (Gomes, Golino, Pinheiro, Miranda, & Soares, 2011). These approaches are a combination of cognitive processing and motivation that drive and qualify interaction. The deep approach implies a cognitive processing that focuses on the formation of meanings, the production of integrated knowledge and understanding and, at the same time, involves a distinctive motivation of the interaction itself. On the other hand, the superficial approach basically implies rote learning and extrinsic motivation (Gomes & Golino, 2012). From the point of view of its social relevance, the theory of approaches has been used for the planning of teaching practices that aim to foster more active and participatory student learning (Contreras, Salgado, Hernández-Pina, & Hernández, 2017). Two meta-analyses (Richardson, Abraham, & Bond, 2012; Watkins, 2001) indicate that the deep approach has a positive correlation with better overall academic performance, both school and university, while the superficial approach has a positive correlation with poorer performance. This evidence supports the theory's argument that deep approach produces a learning of better quality, although the correlations found in these meta-analyses were weak ($r < 0.25$). However, Gomes (2011) provides original evidence that the value of the correlation between approaches and academic achievement may vary across school grades, indicating the likelihood that the deep approach is closely related to performance in particular grades and the superficial approach in other grades. It is possible that meta-analyses, by aggregating school performance in different grades, may find a result that underestimates the magnitude of correlations in specific periods.

In spite of the contributions and advances made, the field of learning approach studies brings up an important gap: the exclusive use of self-report instruments to measure their constructs (Contreras et al., 2017). This is a difficult matter since self-report instruments need people with a good perception or judgment of their own internal processes so that the generated measures do not contain too much noise. In the case of approaches, self-report tools depend on respondents having a good understanding of how they interact with objects of knowledge and know how to identify the behaviors that delimit their interactions and which behaviors are indeed frequent or significant in their personal repertoire. Besides, self-report instruments tend to generate considerable response bias, for example the tendency to choose options which overvalue the score, called social desirability. This is supported by a wide range of research (see Wetzel, Böhnke, & Brown, 2016).

The Students' Learning Approach Test - Thinking Identification in Text (SLAT-Thinking) was recently developed by Cristiano Mauro Assis Gomes and Isabella Santos Linhares, in 2019, from Laboratory for Cognitive Architecture Mapping (Laboratório de Investigação da Arquitetura Cognitiva - LaiCo), aiming to evaluate learning approaches through performance. SLAT-Thinking has two texts and 12 items related to each text. Each item is formed by a statement that may or may not represent the author's thinking in a given text. The respondent's task involves reading each text, reading the statement present in each item, and identifying whether or not that statement represents the author's thinking contained in the text. Figure 1 shows the instruction of SLAT-Thinking. The texts and items are not presented in order to preserve the answer key of the test.

The content of the texts was strategically designed to allow the previous knowledge of the respondent to be intensively activated. Both texts were created by the authors of the test to highlight controversial topics of strong social relevance that are part of the current discussions of modern society. The first theme involves violence against women and the second concerns prejudice against homosexual couples. In theory, controversial topics of strong social relevance contain a rich context for verifying whether the respondent is able to correctly identify the author's thinking structure in a given text. The assumption of each of the two texts has a strong common sense character, referring to well-established prior knowledge, that is, opinions already well formed by most readers. The arguments of each text follow the assumption in order to support it. They do not necessarily oppose common sense values and thoughts about the theme involved in each text, but develop specific ideas of their own to support the assumption. In this context, the reader who does not correctly analyze the logic of the text's arguments, as well as does not have a firm attitude and engagement to accurately identify the author's thinking and differentiate it from

his own or other people's thinking tends to interpret that the text makes a number of statements or have a number of thoughts while in reality this is not true. This difference in strategies and motivations in the use of prior knowledge and analysis of argumentative structure is what marks the deep approach and the superficial approach in identifying the author's thinking in a particular text in SLAT-Thinking.

**Students' Learning Approach Test:
Identification of Thinking Contained in Texts
(SLAT-Thinking)**

Name: _____

Age: _____

Education:_____

This test aims to verify your ability to identify whether or not certain statements represent the author's thinking contained in a given text.

This test consists of two texts. Below each text there is a sequence of items with statements that may or may not represent the author's thinking contained in a given text.

Read the first text and answer the items related to this text. Then, read the second text and answer the items related to this text.

If you think that an item's statement EXPRESSES the author's thinking contained in the text to which the item refers, put an x over column E (express). If you think that the same statement DOES NOT EXPRESS the author's thinking contained in this text, put an x on column N (not express). If you think there is NO ANSWERING if the statement expresses or does not express the author's thinking contained in this text, put an x in column Z (there is no way to answer).

Here's an example:

After reading text 1 (which is not shown in this example) and analyzing the statement in item 1, "Grêmio is the greatest soccer team in the world.", the reader concludes that there is no way to answer whether the statement in item 1 represents or it does not represent the author's thinking contained in text 1. Thus, the reader answers item 1, marking an x in column Z.

STATEMENT	E	N	Z
1) Grêmio is the greatest soccer team in the world.			X

Figure 1. Instruction of the SLAT-Thinking Test.

The SLAT-Thinking is supported theoretically and operationally by some premises and applications, both elaborated by the first author of the test. Due to their importance, they are summarized below.

1. Basic Premise: Learning approaches can be measured through people's performance, taking as observable variables the subject's performance in tasks involving the subject-object interaction of knowledge.

2. Complementary Premise: The approaches are characterized by various behaviors representing cognitive strategies and motivations which qualify the interaction of the subject with the objects of knowledge; build relationships, identify implicit items, construct meaning, are examples of strategies, while the intrinsic and extrinsic motivation to the task are examples of motivations. The measurement of performance approaches imposes the need for each of these behaviors to be measured by different tests, considering that each of these behaviors is best measured by different tasks.

3. Application # 1: SLAT-Thinking evaluates people's approaches to identifying an author's thinking in a given text. This behavior was chosen because it represents a relevant interaction of the subject with the objects of knowledge of the 21st century. The individual who shows the engagement and cognitive strategy of identifying the argumentative structures of his own thinking, differentiating them from the structures of others, is able to understand in greater depth the thought contained in the speeches and texts present in face-to-face interactions, on the internet, in social networks, making him more apt to critically analyze, to build authentic arguments, to criticize responsibly and with reasoning. Moreover, it allows the individual to become less likely to be guided or to sustain his thinking

through fragile, contradictory or superficial arguments. In short, this behavior is a tool for the formation of critical and reflective citizens.

4. Application no. 2: In addition to SLAT-Thinking, other tests may be designed, as this will allow for a wide range of assessments in the future capable of addressing different behaviors that characterize learning approaches, such as building relationships, identifying implicit relationships, transferring knowledge to new contexts.

5. Operational Premise: Prior knowledge and logical analysis of the arguments are defined as the fundamental components for a correct identification of the author's thinking of a given text. Prior knowledge is a basic condition for reading. An argumentative structure is not just a product of logical relations. It expresses meanings, beliefs, opinions, expectations, concepts. Reading an argumentative structure is a conjugation of logical analysis of arguments and identification of word meanings through prior knowledge. Prior knowledge, although mandatory, needs to be well used. An example of misuse is the situation in which the reader reads a certain text and erroneously infers that the author of the text presents certain conceptions, arguments and thoughts, whereas in reality these elements are products of the reader's own prior knowledge. The logical analysis of the arguments, on the other hand, allows the reader to properly identify the argumentative structure of the author of the text. In short, if the logical identification of the argumentative structure depends on the reader's prior knowledge of the topic surrounding the text, logical analysis corrects possible inappropriate uses of prior knowledge and promotes its correct application.

Due to its unprecedented features, SLAT-Thinking has not been evaluated, in terms of evidence about its validity. The present study aims to present the first evidence about the factorial structure of SLAT-Thinking. For that purpose, a heterogeneous sample of higher education students was randomly separated into two groups, a training sample and a test sample. In the training sample two models are tested. The first model, called the continuum model, assumes that a single latent variable explains the variance in people's performance across all test items. This model also assumes the existence of a correlation between two items that deviate from the text theme. This single latent variable represents a continuum between the superficial and deep approach where a higher score on the variable represents a larger deep approach and a lower score represents a larger superficial approach. The second model, called

the level model, assumes that there are four distinct levels of learning approaches: superficial, intermediate-superficial, intermediate-deep, deep. The superficial approach explains items 3, 6, 8, 12, 16 and 24; the intermediate-superficial approach would explain items 2, 11, 22 and 23; the deep intermediate approach explains items 4, 7, 10, 14, 15, 18 and 20; the deep approach would explain items 1, 9, 13, 19 and 21. As the continuum model, the level model assumes that items 5 and 17 correlate. The levels would involve the degree of difficulty of the items and were organized from the observation of the percentage of correct answers of the participants (see Table 2). They would represent superficial or profound strategies of using prior knowledge and logically analyzing the arguments of the text to correctly identify the author's thinking contained in a given text. Unfortunately, it is not possible to detail the processes involved in each item, as this detail would expose the test answer sheet. However, it can be said that the items of the superficial approach of the level model are only right with the application of prior knowledge and common sense about the theme and a superficial analysis of the author's arguments. On the other hand, the items of the intermediate-superficial intermediate approach demand a more refined argumentative analysis, but it is still possible to get them right through the use of previous knowledge and politically correct clichés. Whereas the items of the deep-middle approach require a much more accurate argumentative analysis, and the use of politically incorrect clichés or previous knowledge on the subject may not be sufficient to get them right. Finally, the deep approach items require a very detailed argumentative analysis, as well as a clear separation between the author's arguments contained in the text and the reader's prior knowledge and conceptions. The analyzes in the training sample aim to identify a final model that best represents the factorial structure of SLAT-Thinking. After identification, this final model will be tested in terms of its configural, metric and scalar invariance, comparing the training and test samples. This analysis aims to verify if the model generated in the test sample has generality and is also suitable for the training sample.

Method

Participants

Students from seven higher education institutions in Brazil (N = 622), being 391 (62.9%) from private

institutions and 231 (37.1%) from public institutions were sampled for this study. The sample includes 332 (53.4%) female students and 290 (46.6%) male students, and consists predominantly of young adults, with an average age of 23.7 years (*SD* = 7.1). The students in this sample fall into three broad areas of knowledge: 77 (12.4%) of Biological Science, 257 (41.3%) of exact science and 288 (46.3%) of human science.

Instrument

SLAT-Thinking. The purpose of SLAT-Thinking is to evaluate the subject's learning approach by seeking to identify the thinking of a given author, contained in a given text. The test was created in the Brazilian Portuguese, in 2018, by (to be presented if the article is accepted for publication). The target audience for the test is people with at least incomplete high school.

The respondent's task involves reading two texts of similar size and argumentative structure and answering 12 items related to text 1 and 12 items related to text 2. The respondent should read the first text and answer the items related to that text and then read the second text and answer the items related to it. Each item contains a statement that may or may not represent the author's thinking contained in the text. The respondent should identify whether or not the statement of each item represents the author's thinking contained in the text. Each item has three possible options: E (expresses the author's thought contained in the text), N (does not express the author's thought contained in the text) and Z (there is no way to answer). The respondent must choose one of them when answering each item. The respondent score for each item is composed of 0 (fail) or 1 (pass) defined by an answer key. The maximum allowed test resolution time is 60 minutes.

Collection and Data Analysis Procedures

The students in the sample of this study were selected by convenience and come from two independent collections, both conducted in 2019. One collection includes 513 students and the other 109 students. These collections come from two different projects aiming to build a database on educational variables so that they can be applied to various instruments, including the SLAT-Thinking. The projects were approved by ethics committees of different Brazilian institutions (number of protocols to be presented, if the article is accepted for publication). Both collections comply with the ethical principles and only included students

who agreed to participate in the survey via Informed Consent.

All analyzes were performed on version 3.6.1 of the R software (Core Team, 2019) using the features of versions 0.6-5 and 0.5-1 of the lavaan (Rosseel, 2019) and semTools packages (Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2018), respectively. The sample participants were randomly separated into two groups of similar size: training sample and test sample.

Data analysis involved two fundamental stages. In the first stage, the two models presented in the introduction of this article were tested by confirmatory factor analysis of items. This stage aimed to identify if either model fit the data. Models were not rejected if they had Comparative Adjustment Index (CFI) greater than or equal to 0.90 and Root Mean Square Error of Approximation (RMSEA) less than 0.10 (Putnick & Bornstein, 2016). To compare the models, the hypothesis test of the Chi-square difference and degrees of freedom of the models was used. After the comparison of the models, an exploratory analysis was performed in order to verify whether the model with best adjustment could be improved with the addition of some relationships. The modification index was used to identify the relationships that could be added to this model and improve its adjustment. All of these analyzes were performed on the training sample. The second stage involved testing whether the final model produced in the first stage had generalization features. For this, an invariance analysis of the final model was performed in the test and training samples, testing the configural, metric and scalar invariance model. The configural invariance model assumes that the training and test samples have the same factors. The metric invariance model is more restrictive because it assumes that the samples have the same factors and the same factorial loads. In turn, the scalar invariance model is even more restrictive because it assumes that the samples have the same factors, the same factorial loads and the same thresholds. The first step of this stage involved analyzing whether the configural model would not be rejected, considering the same values as the CFI and RMSEA used in the first stage. The second step involved comparing the metric and scalar models with the configural model in order to verify that these more restrictive models would not be rejected. The most restrictive models were rejected if they presented, in relation to the configural model, a difference greater than 0.002 in the CFI and p value regarding the comparison of Chi-square and degrees of freedom less than 0.01 (Meade, Johnson, & Braddy,

2008). Since the responses of the SLAT-Thinking items are dichotomous (right or wrong), the Weighted Least Square Mean and Variance Adjusted (WLSVM) estimator was selected to perform both factor analysis and invariance analysis. In case the final model proves invariant, the factor loadings, the correlations between the factors as well as the reliability of their scores (Cronbach's Alpha and McDonald's Omega) will be presented in the results, taking the complete sample (training sample and test sample together).

Results

The results of the tested models can be seen in Table 1. The continuum model was rejected because its CFI is below the acceptable value. The level model was not rejected. Furthermore, examination of the comparison of Chi-squares and degrees of freedom indicates that we can reject the null hypothesis that both models have the same fit. The subsequent step of analysis involved checking whether the level model could be improved by adding some relationships Two relations were added, allowing item 9 to be also carried by the deep-intermediate approach factor and item 14 by the deep approach factor, generating the level model (2 relations). Examination of the Chi-square and degree of freedom comparison of the level model and the level model (2 relations) indicates that we can reject the null hypothesis that both models have the same degree of fit.

The results for invariance analysis are also shown in Table 1. As the level model (+2 relations) showed the best fit, it was defined as the final model and used in the invariance analysis. The configuration model was not rejected, indicating that the four approaches of the level model (+2 relations) are present in both training sample and test sample. The metric model was not rejected either, because the null hypothesis test does not allow us to reject the hypothesis that the metric model has the same fit as the configural model. The same applies to the scalar model. This indicates that the level model (+2 relations) is invariant up to the analyzed scalar level, implying evidence of generality of the model.

Table 1. Results of Confirmatory Factor Analysis of Items and Invariance Analysis

Model	χ^2(gl)	$\Delta\chi^2$(Δgl)	p	CFI	RMSEA	ΔCFI	IC 90%
Fatorial analysis							
Continuum[*5]	474.91(251)			0.869	0.054		0.047-0.062
Level model[*5]	388.49(247)	31.62(4)[*1]	.000002	0.917	0.044		0.035-0.052
Level model (+2 relations)[*5]	349.68(245)	16.42(2)[*2]	.000272	0.939	0.038		0.028-0.046
Invariance analysis							
Configural[*6]	757.20(490)			0.934		0.042	0.036-0.048
Metric [*6]	799.73(510)	24.36(20)[*3]	.2272	0.929	0.043	0.005[*3]	0.037-0.048
Scale [*6]	789.25(506)	22.56(16)[*4]	.1261	0.930	0.042	0.004[*4]	0.037-0.048
Factorial analysis							
Level Model (+2 relations)[*7]	447.93(245)			0.946	0.037		0.031-0.042

Note. χ^2 = chi-square; gl = degrees of freedom; CI = confidence interval; * 1 = comparison between the level model and the continuum model; * 2 = comparison between level model and level model (+2 relations); * 3 = comparison between metric and configural model; * 4 = comparison between scalar and configural models; * 5 = analysis in the training sample; * 6 = invariance analysis of the training sample versus test ; * 7 = analysis in the complete sample.

Figure 2 shows the model relations of level (+2 relations) in the complete sample. The factor loadings are standardized. The superficial approach factor shows an average factor load of 0.66 (SD = 0.20), a minimum of 0.34 and a maximum of 0.86, all showing p-value <0.0009; the superficial-intermediate approach factor exhibits an average load of 0.34 (SD = 0.06), a minimum of 0.25 and a maximum of 0.39, all showing p-value <0.025; the deep-intermediate approach factor has a mean load of 0.42 (SD = 0.10), a minimum of 0.29 and a maximum of 0.55, all with p-value <0.002; The deep approach factor has a mean load of 0.50 (SD = 0.21), a minimum of 0.24 and a maximum of 0.75, all with p-value <0.01.
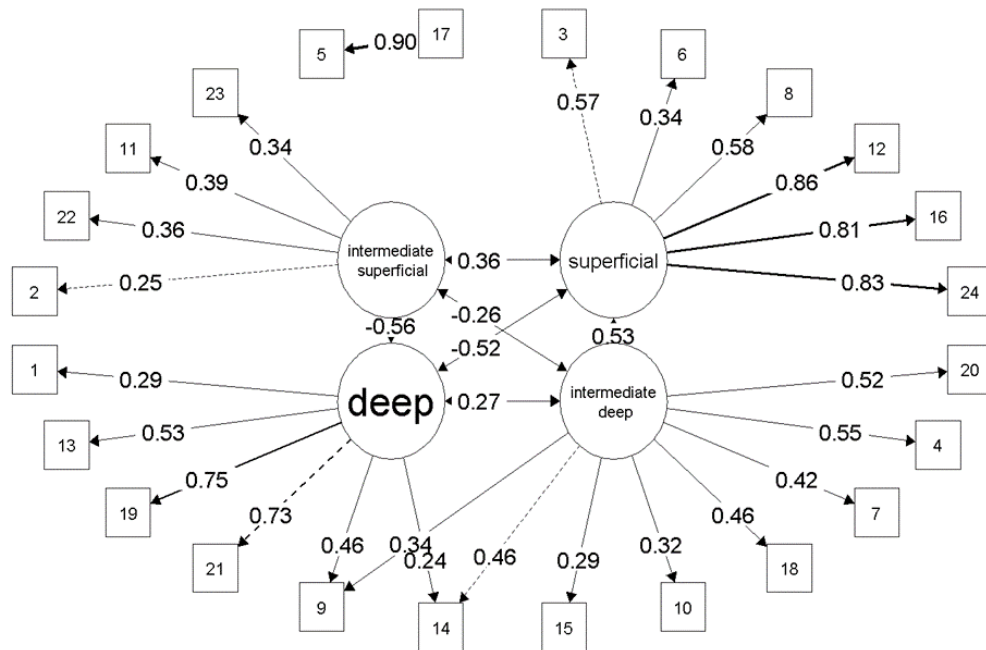
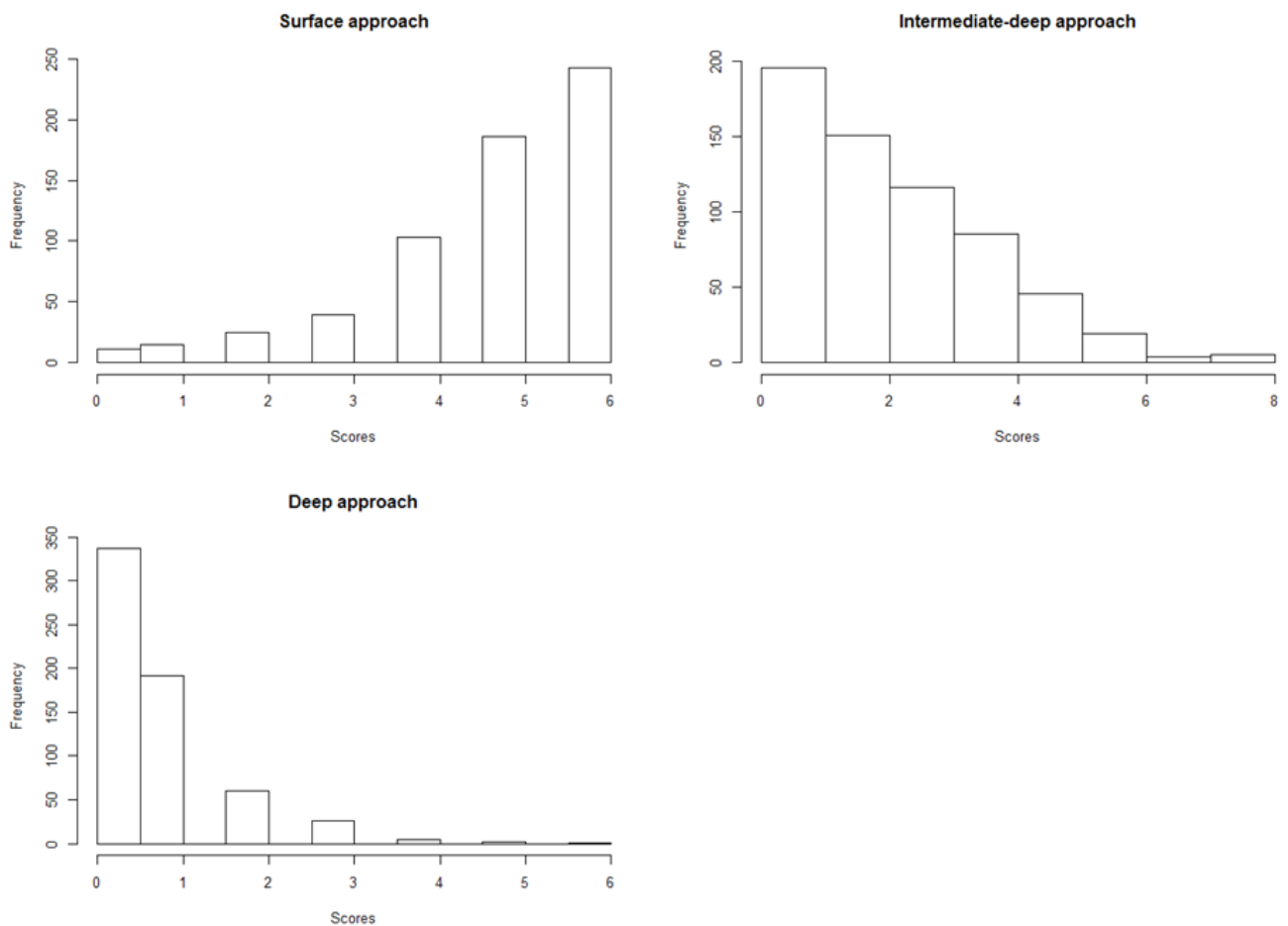**Figure 2.** Factorial Structure of the Level Model (+ 2relations).

Figure 3. Histogram of the Distribution of the Raw Scores of the Approaches.

The correlations between the factors have *p*-value <0.04, except for the correlation between the intermediate-superficial and deep-intermediate approach factors (*p*-value = 0.09), so this correlation is being interpreted as null. There is a negative correlation of the deep approach with the superficial and intermediate-superficial approaches. The deep-intermediate approach has a positive correlation with the deep and superficial approaches. The factors of the superficial (α = 0.82 and ω = 0.61), intermediate-deep (α = 0.66 and ω = 0.41) and deep (α = 0.69 and ω = 0.51) approaches exhibited acceptable reliability, considering the alpha index of Cronbach. The intermediate-superficial approach factor did not produce reliable scores (α = 0.34 and ω = 0.23).

Figure 3 shows graphically the distribution of gross scores on the factors of the approaches. The superficial-intermediate approach factor was not included because it did not have acceptable reliability. Regarding the six superficial approach items, 14% (*N* = 90) of the participants hit up to three items, while 86% (*N* = 532) hit four to six items. In the case of the eight intermediate-deep approach items, 74% (*N* = 463) hit up to three items and 26% (*N* = 159) hit four to eight items. Finally, the six deep approach items had 85% (*N* = 529) hit none or only one item, while 14% (*N* = 86) hit two or three items, and only 1% (*N* = 7) hit four to six items. Table 2 shows the difficulty parameters of the SLAT-Thinking items.

Table 2. Difficulty Parameters of the SLAT-Thinking' Items

Items	Correct answers	Thresholds	p-value
1	10%	1.302	0.000
2	43%	0.178	0.000
3	77%	-0.744	0.000
4	23%	0.729	0.000
5	45%	0.137	0.006
6	68%	-0.472	0.000
7	47%	0.081	0.109
8	73%	-0.622	0.000
9	11%	1.247	0.000
10	42%	0.199	0.000
11	66%	-0.415	0.000
12	88%	-1.188	0.000
13	09%	1.341	0.000
14	27%	0.617	0.000
15	21%	0.804	0.000
16	85%	-1.025	0.000
17	43%	0.186	0.000
18	47%	0.085	0.092
19	04%	1.730	0.000
20	28%	0.574	0.000
21	08%	1.436	0.000
22	59%	-0.240	0.000
23	58%	-0.203	0.000
24	88%	-1.164	0.000

Note: Thresholds were standardized in mean 0 and standard deviation 1.

Discussion

The rejection of the adjustment of the continuum model indicates that SLAT-Thinking does not measure a continuum where one pole of it represents the superficial approach and the other pole the deep approach. On the other hand, the acceptable adjustment of the level model indicates that SLAT-Thinking measures

different levels of learning approaches and that these levels correlate. This evidence becomes more robust in the invariance analysis, since the four latent variables of the level model, their factor loadings and thresholds are the same in the training and test sample.

Despite the levels of model indicating the validity of four levels of approaches, SLAT-Thinking evaluates reliably the superficial approaches, intermediate-deep and deep, as the factor scores of these levels showed a higher than 0.60 point alpha, which is the lowest cutoff to be considered an acceptable score. The factorial score of the intermediate-superficial approach however proved unacceptable, and should not be used. New items related to this level should be elaborated and incorporated into the test so that a minimally reliable score for the intermediate-superficial approach can be achieved in the future. In addition to alpha, we also present the omega index for the level model factor scores. To date, we have no well-defined reference in the psychometric literature as to what should be the cutoff points for the omega index. We know that it is much more demanding than the alpha index because it considers the factor loadings of the items in its calculation, unlike the alpha that does not use this information (Gomes, Golino, & Peres, 2018; Valentini et al., 2015). Considering this issue, we prefer only to report the obtained omega, but not to produce any inference about it, until the international psychometric community considers possible omega cutoff points as an indicator of the reliability of the factor scores.

The negative correlation of the deep approach with the superficial approaches, as well as the positive correlation of the deep intermediate approach with the deep and superficial approaches seem to indicate that the deep intermediate approach is a kind of bridge/connection that links the deep and superficial approaches. The negative correlation suggests that the deep approach is incompatible or opposite to the surface approaches. However, the intermediate-deep approach, from its positive correlation with the deep and superficial approaches, is compatible with both. If the superficial approach indicates a way of identifying the author's thought through clichés and a shallow logical analysis, the deep approach implies its opposite, that is, an identification of the author's thinking through an accurate logical analysis and use of prior knowledge without mixing the reader's thinking with the author's thinking. In this sense, the negative correlation found between these approaches makes theoretical sense. In turn, the positive correlation of the deep-intermediate approach with these approaches indicates that possibly the deep-intermediate approach involves a more thorough logical analysis than the superficial one, but still has certain

characteristics of a not yet fully deep analysis. By merging some of the superficial and deep approach, it makes theoretical sense that the intermediate-deep approach correlates positively with both.

The implications of the evidence found in the factorial structure and invariance of SLAT-Thinking are relevant to the field of approach measures as well as for studies in educational psychology. This paper presents the first evidence that approaches can be conceived not in terms of "all or nothing", but in terms of transitions. The intermediate-superficial approach and the deep-intermediate approach bring this new perspective to the theory of learning approaches. To date, the field has only dealt with the empirical identification of the superficial approach, the deep approach, and the strategic approach (Contreras et al. 2017). This new perspective is interesting since it brings new possibilities of study in the area, particularly with regard to the possibility of developing the deep approach from the superficial approach, fostering, in this way, the boost of intermediate levels. This field of research is an unknown area that deserves to be explored, as it has the potential to generate knowledge about both the development of the shallow and deep approach and the possible interventions aimed at promoting deeper approaches that can foster a learning of better quality.

It is important to take up the argument that SLAT-Thinking does not evaluate all behaviors linked to the manifestation of the learning approach. The postulates that support SLAT-Thinking are clear in stating that no performance test could possibly assess the diversity of behaviors that make up learning approaches. In short, the SLAT-Thinking evaluates the learning approaches of people to identify the thought of an author, contained in a given text. However, this behavior is the key to the formation of the critical citizen of the 21st century. Properly understanding the argumentative structures present in speeches, texts, the internet, political speeches, etc., and knowing how to discern the thought contained in these structures, differentiating it from his or hers own thought is a fundamental action for the citizen to have a lucid judgment and critic about the different worldviews, opinions, ideological perspectives, among other elements that lead to public life. Therefore, the learning approach assessed by SLAT-Thinking is relevant and can be used for diagnostic purposes of teaching quality. For example, the National High School Examination (ENEM) aims to evaluate the education of the Brazilian student who finishes high school, from the perspective of fostering a critical student, able to analyze arguments, reason in depth, and discern ideas and thoughts with clarity (Gomes, 2010a; Gomes & Borges,

2009). In this sense, SLAT-Thinking is an excellent tool for assessing whether ENEM has fulfilled its promises. This can be seen by verifying whether performance on ENEM items involving text interpretation and critical thinking correlates with more superficial or more advanced levels of people's approach to identify an author's thinking in a given text. We give the example of ENEM because this Brazilian large-scale evaluation is very important. Nonetheless we may also foresee the SLAT-Thinking as a tool to investigate whether evaluations and academic practices have led students to read and analyze arguments supported by clichés, politically correct stereotypes, previous knowledge that is confused with the author's thinking or supported by an accurate logical argumentative analysis and a precise use of the previous knowledge to understand the elements, without being confused with the argumentative structure of the analyzed text. In short, SLAT-Thinking allows, even indirectly, to assess whether or not academic assessments and practices are enabling the generation of critical citizens.

The sample used in this study is not small, nonetheless not very large and was composed by a great diversity of university students from different public and private institutions. In this sense, the sample of this study is heterogeneous in terms of academic performance. Therefore, it is relevant to point out that most of the people of the sample presented a large number of success, correct answers, in the superficial approach items, that is, the participants of this study demonstrated to know how to identify the author's thought when the process demands a shallow argumentative analysis and the use of politically correct clichés or stereotypes. However, the participants showed a very small number of correct answers in the deep approach, indicating poor argumentative analysis and discernment between their own prior knowledge and the author's thinking contained in a given text. Considering the heterogeneity of the sample, it is possible that this pattern may be found in new studies and may represent a trend in Brazilian university students. Clearly further studies need to be conducted so that this hypothesis can be known, but it is possible that the results of this study bring evidence that the Brazilian education promotes overall, a thought of poor quality. This result is in line with various evidences that Brazilian learning in general is of low quality. To cite one example, the Paulo Montenegro Institute provides solid evidence that about 12% of Brazilians aged 15-64 have proficient functional literacy. That is, 88% of Brazilians in this age group do not demonstrate proficiency in reading texts, interpreting and producing inferences from them. Moreover, the Paulo Montenegro Institute provides evidence that this percentage has remained relatively stable from the early 2000s until the present time (Instituto Paulo Montenegro, 2018). The report also provides striking evidence that only 34% of people in this age group in tertiary education in 2018 had proficient functional literacy. This is serious, since the rapid growth of access to Brazilian higher education by economically underprivileged groups occurred precisely during this period, so that some important change in the reading ability of Brazilians between 15 and 64 years should have occurred.

Conclusion

This article brings some contributions to the literature. The first one involved presenting the first evidence regarding the structural validity of the first test that measures learning approaches through performance. The second contribution brought new perspectives on the learning approaches construct. The original evidence found that approaches may not be all-or-nothing, involving intermediate dimensions, opens a new research agenda regarding the development of approaches and their implications for teaching and learning. The third contribution involved showing how SLAT-Thinking can be a relevant tool for the diagnosis of educational quality, in terms of critical citizen formation.

SLAT-Thinking brings two practical implications. While the existing instruments of learning approaches measure perceptions of students about their own approaches, SLAT-Thinking do not measure perceptions, but learning approaches itself in a specific domain. Furthermore, SLAT-Thinking does not measure approaches of learning in an all-or-nothing paradigm, as the one in which a student has just a superficial or deep approach. Instead of that, SLAT-Thinking identifies different levels of learning approaches, possibly implying stages of development of learning approaches.

However heterogeneous the sample employed in this study has been, it is relevant that further studies be carried out, assessing the factorial structure of the SLAT-Thinking, in order to generate increasingly robust evidence regarding the approach level model. We highlight the difficulty of finding participants with a high profound approach. In addition, it must be considered that structural validity is one of several analyzes of the validity of an instrument. SLAT-Thinking needs to be investigated in terms of its external validity, both in terms of its predictive ability in relation to academic performance and its association with related constructs that are also based on the perspective of an active subject in its interaction with the objects of knowledge.

References

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. doi: 10.5294/edu.2017.20.1.4

Core Team. (2019). R (version 3.5) (Computer software). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

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

Gomes, C. M. A. (2010b). Perfis de estudantes e a relação entre abordagens. *Psico, 41*(4), 503-509. Retrieved from http://revistaseletronicas.pucrs.br/revistapsico/ojs/index.php/revistapsico/issue/view/453

Gomes, C. M. A. (2011). Abordagem profunda e abordagem superficial à aprendizagem: diferentes perspectivas do rendimento escolar. *Psicologia Reflexão & Crítica, 24*(3), 438-447. doi: 10.1590/S0102-79722011000300004

Gomes, C. M. A. (2013). A construção de uma medida em abordagens de aprendizagem. *Psico, 4*(2), 193-203. Retrieved from http://revistaseletronicas.pucrs.br/revistapsico/ojs/index.php/revistapsico/issue/view/664

Gomes, C. M. A., & Borges, O. N. (2009). O ENEM é uma avaliação educacional construtivista? Um estudo de validade de construto. *Estudos em Avaliação Educacional, 20*(42), 73-88. doi: 10.18222/eae204220092060

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

Gomes, C. M. A. Golino, H. F., & Peres, A. J. S. (2018). Análise da fidedignidade composta dos escores do Enem por meio da análise fatorial de itens. *European Journal of Education Studies, 5*(8), 331-344. doi: 10.5281/zenodo.2527904

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

Instituto Paulo Montenegro. (2018). *Indicador de Analfabetismo Funcional*. Retrieved from https://drive.google.com/file/d/1ez-6jrlrRRUm9JJ3MkwxEUffltjCTEl6/view

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). *semTools: Useful tools for structural equation modeling. R package* (version 0.5-1) (Computer software). Retrieved from https://CRAN.R-project.org/package=semTools

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. doi: 10.1037/0021-9010.93.3.568

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71-90. doi: 10.1016/j.dr.2016.06.004

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi: 10.1037/a0026838

Rosseel, Y., Jorgensen, T. D., Oberski, D., Vanbrabant, J. B. L., Savalei, V., Hallquist, E. M., … M., Scharf, F. (2019). *lavaan: Latent Variable Analysis. R package* (version 0.6-5) (Computer software). Retrieved from https://cran.r-project.org/web/packages/lavaan/index.html

Valentini, F., Gomes, C. M. A., Muniz, M., Mecca, T. P., Laros, J. A., & Andrade, J. M. (2015). Confiabilidade dos índices fatoriais da Wais-III adaptada para a população brasileira. *Psicologia: Teoria e Prática, 17*(2), 123-139. doi: 10.15348/1980-6906/psicologia.v17n2p123-139

Watkins, D. (2001). Correlates of Approaches to Learning: A Cross-Cultural Meta-Analysis. In R. J. Sternberg & L. F. Zhang (Eds.), *Perspectives on thinking, learning and cognitive styles* (pp. 132-157). Mahwah, NJ: Lawrence Erlbaum Associates.

Wetzel, E., Böhnke, J. R., & Brown, A. (2016). Response Biases. In F. T. L. Leong, D. Bartram, F. M. Cheung, K. F. Geisinger, & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (pp. 349-363). New York, NY: Oxford University Press.

Cristiano Mauro Assis Gomes, Doutor em Educação pela Universidade Federal de Minas Gerais/UFMG, é Professor da Universidade Federal de Minas Gerais/ UFMG e Coordenador do Laboratório de Investigação da Arquitetura Cognitiva - LAICO/UFMG. Endereço para correspondência: Avenida Antônio Carlos 6627, Bairro Pampulha, Belo Horizonte, Minas Gerais. CEP: 31.270-010. Email: cristianomaurogomes@gmail.com

Josiany Soares Quadros, Mestre em Neurociências pela Universidade Federal de Minas Gerais/UFMG, é Professora da Faculdade Pitágoras. Email: josiany.quadros@atosrh.com.br

Jhonys de Araujo, Mestre em Neurociências pela Universidade Federal de Minas Gerais/UFMG, é Membro do Laboratório de Investigação da Arquitetura Cognitiva - LAICO/UFMG. Email: jhonys.bio@gmail.com

Enio Galinkin Jelihovschi, Doutor em Estatística pela Universidade da Califórnia, Berkeley, EUA, é Professor da Universidade Estadual de Santa Cruz, Ilhéus, Bahia. Email: eniojelihovs@gmail.com