Value-added in Higher Education:
a new methodology for undergraduate programs in Brazil

Valor-agregado no Ensino Superior:
uma nova metodologia para os cursos de graduação no Brasil

Vivian Duarte Couto Fernandes [1]

Gilberto José Miranda [2]

Nicola Alexander [3]

Janser Moura Pereira [4]

Abstract: Among the quality indicators released by the Brazilian Higher Education Assessment System (Sinaes), the Indicator "Difference between Observed and Expected Performance" (IDD) has the purpose of measuring the contribution of the course to student achievement during undergraduate programs. The research presented here offers a new methodology for calculating the IDD (Model IDD-VDCF), examining the philosophical and statistical underpinnings of quality measures, focusing on those that capture the value-added as a student achievement growth. The survey included a sample of 30,668 students, from 911 accounting undergraduate programs in Brazil. The insertion of control variables (at the student and at the institution level) reduced the bias of the IDD estimate associated with the student's selection in specific Accounting Sciences courses. The results call attention to the need to consider the students' learning context when one wants to compare the performance between institutions based on standardized tests. The major contribution of this work is the development of a measure that disentangles more fully what the contribution of program is to student learning, and what merely is a reflection of the capacity that a student brought to the program.

Keywords: Value-added model. Learning outcomes. Quality measures.

Resumo: Dentre os indicadores de qualidade divulgados pelo Sistema Nacional de Avaliação do Ensino Superior (Sinaes) no Brasil, o Indicador de Diferença entre os Desempenhos Observado e Esperado (IDD) tem a proposta de mensurar a contribuição do curso para o desempenho do estudante durante a sua graduação. Esta pesquisa oferece uma nova metodologia de cálculo para o IDD (Modelo IDD-VDCF), examinando os fundamentos filosóficos e estatísticos das medidas de qualidade, com foco naquelas que capturam o valor agregado como um crescimento do conhecimento do estudante. A pesquisa incluiu uma amostra de 30.668 estudantes provenientes de 911 cursos de graduação em Ciências Contábeis no Brasil. A inserção das variáveis de controle (ao nível do estudante e da instituição) reduziu o viés da estimação do IDD associado à seleção do estudante em cursos específicos de Ciências Contábeis. Os resultados chamam atenção para a necessidade de se considerar o contexto de aprendizagem dos estudantes quando se quer comparar o desempenho entre instituições com base em testes padronizados. A maior contribuição deste trabalho é o desenvolvimento de uma medida que esclarece melhor qual é a contribuição do curso para o desempenho do estudante e o que é apenas um reflexo da capacidade prévia deste estudante.

Palavras-chave: Modelos de valor agregado. Resultados de Aprendizagem. Medidas de qualidade.

[1] Universidade Federal de Uberlândia | Faculdade de Gestão e Negócios | Uberlândia | MG | Brasil. Contato: vivian@ufu.br. ORCID: https://orcid.org/0000-0003-4046-5577

[2] Universidade Federal de Uberlândia| Programa de Pós-Graduação em Ciências Contábeis | Uberlândia | MG | Brasil. Contato: gilbertojm@ufu.br. ORCID: https://orcid.org/0000-0002-1543-611X

[3] University of Minnesota | College of Education and Human Development | Minneapolis| MN | EUA. Contato: nalexand@umn.edu . ORCID: https://orcid.org/0000-0002-4650-6563

[4] Universidade Federal de Uberlândia | Faculdade de Matemática | Uberlândia | MG | Brasil. Contato: janser@ufu.br. ORCID: https://orcid.org/0000-0002-4622-6203

1 Introduction

In Brazil the National Assessment System for Higher Education (Sinaes) was created through Act No. 10861, as of 2004, to assess the quality of higher education institutions (HEI), their undergraduate programs, and their students' academic achievement. It targets various dimensions of education, including teaching, research, extension and outreach, social responsibility, program coordination, faculty, and facilities (BRASIL, 2004). One of the Sinaes quality indicators is the Indicator of Difference between Observed and Expected Achievements (IDD) that measures the value that an undergraduate program adds to the development of its seniors by probing their achievements on the Enade[1] as compared to their developmental characteristics at the beginning of their study track" (INEP, 2017, p. 1).

In higher education, value-added can be defined as the difference between the college seniors' achievement and that of the freshmen, which gives an estimate of how much a student has learnt in a given period (LIU, 2011a). Approaches that attempt to identify value-added dimensions provide clearer insights into what has been transformed, but their downside is that they require a representative measure of the outputs. In Brazil the National Exam of Student Achievement (Enade) is used to assess the achievement of those graduating from undergraduate programs. Applied by the National Institute of Educational Research Anísio Teixeira (Inep) since 2004, it is a standardized knowledge exam that assesses: 1) the students' development against the expected syllabus defined by the national curriculum guidelines for the undergraduate studies in a given domain, 2) their development of competences and skills necessary for solid professional practice, and 3) their awareness of the current state of affairs in Brazil and worldwide (INEP, 2018).

The literature suggests that different methods of estimating academic gain produce different findings (FERNANDES; MIRANDA; ALEXANDER, 2020; KIM; LALANCETTE. 2013; LIU, 2011a; MELGUIZO *et al*., 2017; PIKE, 2016; STEEDLE, 2012). Like any other assessment model, the value-added model (VAM) cannot be a standalone parameter to underlie or bear out public policies, and its modelling require caution in fitting school and family characteristics, so as not to reinforce the disadvantages of HEI which have a relatively low percentage of students from lower socioeconomic background. Yet, it is still possible to use the VAM findings to compare the units under scrutiny by looking into the institutions' achievements against the mean, which includes all other institutions (LIU, 2011a; NATIONAL RESEARCH COUNCIL, 2010).

---

1 Enade (*Exame Nacional de Desempenho dos Estudantes* - National Exam of Student Achievement) is an exam for measuring the achievement of higher education students in specific study tracks.

However, the IDD, which in theory should measure a program's contribution to the seniors' academic achievement, does not include the significant variables that do predict achievement. While still appreciating the efforts made within the Sinaes to identify and measure contribution of the undergraduate programs to student achievement through the IDD, this research aims to come up with a new proposal of IDD (IDD-VDCF Model) by not only focusing on determinants of achievement but also introducing them into a new recommendation of a value-added estimation in undergraduate programs.

2 The Sinaes quality indicators

Data provided by the Higher Education Census lay bare the complexity and importance of this level of education in Brazil, given the significant number of enrollments, undergraduate programs and institutions across the country. The numbers for undergraduate studies have increased significantly over the last ten years: overall it has increased 51% for new students, 56.4% for enrollments, and 52% for graduates (INEP, 2019). Undergraduate programs in the fields of Business, Law and Social Sciences are the most demanded programs in Brazil, representing 30.9% of all enrollments in higher education.

Given such significant growth, the assessment of education quality is relevant not only for national authorities, who need to follow up on the outcomes and accountability of their policies, but also for other stakeholders, including the society concerned with its own academic education. In this context, the Sinaes was created to ensure a national process for assessing HEI, undergraduate programs, and student achievement (BRASIL, 2004).

Two of the quality indicators deserve especial attention for the purposes of this article: the IDD and the Preliminary Program Quality Level (CPC). The IDD is the difference between the students' observed Enade scores and predicted scores based on their admission scores – to estimate the achievements resulting from the students' characteristics before higher education, the estimation includes their scores on the Enem, an exam taken by the end of high school and widely used for admission to the Brazilian undergraduate programs. In other words, this indicator aims to quantify how much each higher education institution adds to their students' achievements during their undergraduate studies. Its relevance for Program Quality Level is borne out through the 35% share that it contributes to the CPC, as it is presented in methodology section.

In several countries, institutional effectiveness is assessed primarily by student achievement measures, usually provided as mean scores in standardized tests (MILLA; SAN MARTÍN; BELLEGEM, 2016), percentage of graduates (BAILEY; XU, 2012) or

achievement growth using value-added measures (MELGUIZO *et al*., 2017; SHAVELSON *et al*., 2016).

3 Value-Added Models (VAM) in Higher Education

In K-12 education, the VAM have been criticized for its use in support policy in evaluating teachers. Alexander, Jang, and Kankane (2017) as well as Alexander and Jang (2019) found that including VAM in teacher evaluation models was not an effective way of improving student achievement in the US. It had negligible improvements in reading and no significant improvement in math. Also, it did not result in reduced disparities among key student groups.

On the other hand, VAM is used in HE to measure student learning gain and answer questions such as: "What is the proportion of variance in student achievement that can be ascribed to schools?", "How effective is a school in achieving results?", or "What institutional features or practices are associated with effective schools?" (KIM; LALANCETTE, 2013, p. 5).

Recent studies have estimated the value-added of HEI in different countries by using regression equations that include independent variables related to students and institutions (i.e., variables that were not linked to their policy, but rather to uncontrollable factors) (BOGOYA; BOGOYA, 2013; CUNHA; MILLER, 2014; KIM; LALANCETTE, 2013; LIU, 2011a, 2011b; MELGUIZO *et al*., 2017; MILLA; SAN MARTÍN; BELLEGEM, 2016; PIKE, 2016; SHAVELSON *et al*., 2016; STEEDLE, 2012).

The different VAM share the fact that they all use knowledge tests as a measure of prior achievement. They also fit variables related to student characteristics or to school context, but studies have no consensus on which variables to include. All models eventually show that some schools are significantly better or worse than the mean. Since models differ in how they use data (years, assumptions, missing data, and variable fitting), their results are not the same (FERNANDES; MIRANDA; ALEXANDER, 2020; NATIONAL RESEARCH COUNCIL, 2010).

Melguizo *et al*. (2017) used one single database to compare three VAM: 1) fixed effects, 2) random effects, and 3) aggregated residuals effects. Building on data from Colombia, the authors employed as response variables the graduation rate, the employment rate of graduates, and the SABER PRO[2] score. The ranking of the HEI changed for each

---

[2] SABER PRO is an exam similar to ENEM in Brazil. The students usually take them in high school.

response variable, which suggests that different achievement variables should be explored to assess educational quality. The study showed empirical evidence to support that the VAM based on fixed effects of the HEIs was the best alternative to address student selection bias.

The very existence of different methodologies of value-added estimation evinces unresolved issues in the literature. Three controversial dimensions stand out in debates amongst researchers and practitioners, namely: 1) use of the VAM and its possible consequences in higher education, 2) VAM measurement methods, and 3) VAM statistics. Criticisms in the first dimension can be reduced by: 1) making explicit the methodological choices and the purpose of estimation, which eventually allows for comparisons and identification of trade-offs; 2) using other achievement metrics to make the model reliable and valid for important policy decisions; and 3) predicting the consequences of the potential incentives that the VAM may awaken in the stakeholders in the educational process (NATIONAL RESEARCH COUNCIL, 2010).

Based on the assumptions of the value-added approach, the conceptual framework underling model design is focused on variables that usually escape institutional control but are predictive of academic achievement. Consequently, the VAM include demographic and contextual factors, which are less susceptible to institutional control. In contrast, they omit variables that represent policies subject to institutional action, because such policies are directly related to the institution's own effectiveness.

4 Methods

4.1 Practical decisions

Shavelson *et al*. (2016) suggest that because it is very likely that educational experience at different colleges (e.g. engineering, sciences, liberal arts) is not the same, the treatment chosen should be declared. Accounting has been one of the five largest fields of undergraduate studies in Brazil since 2009. It currently ranks third in number of undergraduate programs in the country and fourth in number of enrollments, only falling below Law, Pedagogy, and Business (INEP, 2019). Yet, the numbers are disturbing in other fronts. Only 30% (11,210 out of the 37,051) applicants passed the 2019-2 proficiency exam applied by the Federal Board of Accountants. Meanwhile, 1,101 undergraduate programs in Accounting had students taking the 2019 Enade and obtained the following quality levels: 50 rated 1 (worst level); 348, 2; 478, 3 (satisfactory level); 166, 4; and 42, 5 (best level) (INEP, 2019), i.e., 36% of them did not reach a satisfactory quality level according to the criteria set forth by Sinaes.

About the unit choice, the main database was arranged at the student level and the analysis performed at two levels: that of the student, and that of the institution (the undergraduate program in Accounting). The outcomes decision came out from Sinaes. In Brazil, all senior students are required to get the Enade exam which has been applied by Inep since 2004. The present study assumes that Enade is a solid, representative exam of the academic achievement provided by undergraduate programs in Brazil.

4.2 Data-set

It was used longitudinal data from senior students who took Enade in 2015. Prior achievement was measured by Enem score. The study is limited to undergraduate programs in Accounting in Brazil and based on public 2015 databases made available by Inep, namely: Enade microdata, CPC microdata, and IDD microdata. After concatenating all databases and considering that 65,283 Accounting undergraduates had signed up for the 2015 Enade, the sampling resulted in a total of 30,668 students from 911 programs, representing 46.98% of the population.

4.3 Model IDD-Inep

In order to compare the model proposed with the current methodology that calculates the value-added in higher education used by Brazilian government (IDD-Inep), it is important to describe both of them. So, in Inep model, a Hierarchical Linear model is used to estimate the IDD as follows:

$$IDD_{ij} = C_{ij} - \hat{I}_{ij} \tag{1}$$

where: the IDDij is the estimate of the part of the achievement of student i resulting from the quality of the learning conditions provided by the undergraduate program j.

The IDD estimation employs two-level hierarchical linear modelling. One level is that of the student, estimated through:

$$C_{ij} = \beta_{0j} + \beta_{1j} * CN_{ij} + \beta_{2j} * CH_{ij} + \beta_{3j} * LC_{ij} + \beta_{4j} * MT_{ij} + \lambda_{ij} \tag{2}$$

where: C_{ij} is an achievement estimate for senior student i on the Enade as weighted by his/her scores in the specific domain section (75%) and in the general domain section (25%) for the undergraduate program j; CN_{ij} is the measure of achievement in Enem section 'Natural sciences and their technologies' for senior student i in undergraduate program j; CH_{ij} is the measure of achievement in Enem section 'Humanities and their technologies' for senior student i in undergraduate program j; LC_{ij} is the measure of achievement in Enem section

"Languages, codes, and their technologies" for senior student i in undergraduate program j; MT_{ij} is the measure of achievement in Enem section "Mathematics and its technologies" for senior student i in undergraduate program j; λ_{ij} is the random effects associated with senior student i in undergraduate program j.

The second level of analysis is the program, as estimated through:

$$\beta_{0j} = \beta_{00} + u_{oj} \tag{3}$$

where: β_{00} represents the mean or general intercept, which is constant across the programs; and u_{oj} is the random effects associated with undergraduate program j.

The multilevel regression model is estimated twice. The first regression extracts the parameters, estimates the standardized residual and excludes those with a modular value higher than 3. The second regression uses the parameter values to produce the estimate Î as in:

$$\hat{I}_{ij} = \hat{\beta}_{0j} + \hat{\beta}_{1j} * CN_{ij} + \hat{\beta}_{2j} * CH_{ij} + \hat{\beta}_{3j} * LC_{ij} + \hat{\beta}_{4j} * MT_{ij} \tag{4}$$

where: $\hat{I}_{ij}$ is the estimate of the part of the Enade achievement of senior student i in program *j* resulting from the students' characteristics before admission to the program.

A gross IDD_{ij} is estimated for each student i from undergraduate program j, as in equation 1; then, a mean IDD_{ij} is estimated for each program (sum of all IDD_{ij} for program j divided by the number of students from program j). As with the other variables that make up the CPC indicator, the IDD_j score is standardized and transformed into a continuous scale from 1 to 5.

In turn, the CPC is a weighted sum of means related to student achievement (i.e., IDD and Enade score), faculty characteristics, and the program structure in the students' perception, as shown in Equation 5:

$$CPC_c = 0{,}2 * Enade_c + 0{,}35 * IDD_c + 0{,}075 * Me_c + 0{,}15 * Doc_c + 0{,}075 *$$
$$RT_c + 0{,}075 * ODP_c + 0{,}05 * IFF_c + 0{,}025 * OAF_c \tag{5}$$

where CPC_c is the score of the Preliminary Program Level for undergraduate program c; $Enade_c$ is the Enade scores of seniors in undergraduate program c; IDD_c is the IDD score for undergraduate program c; Me_c is the score for the ratio of faculty members with a master's degree in undergraduate program c; Doc_c is the score for the ratio of faculty members with a doctoral degree in undergraduate program c; RT_c is the score for the types of employment contract in undergraduate program c; ODP_c is the score for pedagogical teaching structure (ODP – *organização didático-pedagógica*) in undergraduate program c; IFF_c is the score for infrastructure and physical facilities (IFF – *infraestrutura e instalações físicas*) in

undergraduate program c; OAF_c is the score for opportunity for further learning (OAF – *oportunidades de ampliação da formação*) in undergraduate program c.

4.4 Model IDD-VDCF

Model IDD-VDCF is a value-added model that can be used to estimate how much higher education institutions contribute to the student's final achievement, considering their sociodemographic characteristics, their prior achievement, and the specific characteristics of their HEI. This is a multilevel mixed effects model (random intercept for the program, and fixed effects of the explanatory variables), where value-added is the mean difference between observed scores and estimated scores for all students in a given program.

Thus, the value-added calculated by the IDD-VDCF model is given by:

$$IDD\_VDCF_{ij} = Enade_{ij} - \widehat{Enade}_{ij} \qquad (6)$$

where, $Enade_{ij}$ is the performance measure of the senior accounting student *i* in Enade, and $\widehat{Enade}_{ij}$ is the predicted value of his/her performance considering his/her personal characteristics and the context of the course offer.

Using a database from the Enade editions for the years 2006, 2009, 2012 and 2015 Fernandes *et al.* (2018) identified that the following variables were unanimous (significant) to explain the achievement in all Enade editions analyzed: gender, marital status, income, number of books read per year, hours of extra-class weekly study, participation in academic activity, type of academic organization, region of course offer, infrastructure, pedagogical teaching structure. Therefore, IDD-VDCF model included the significant determinants of achievement in all Enade editions as explanatory variables.

However, the variables related to institutional policies, such as conditions of learning (infrastructure and pedagogical teaching structure), were not included because they were deemed as an integral part of the very program management policy. Hence, including such variables would undermine the actual value-added of the programs. Similarly, the variable related to students' engagement in teacher assistance, research and extension was not included because the existence of such activities was also considered to be part of the institutional policy of the undergraduate program. Thus, the equation that describes Model IDD-VDCF is given by:

Level 1:
$$Enade_{ij} = \beta_{0j} + \beta_1 CN_i + \beta_2 CH_i + \beta_3 LT_i + \beta_4 MT_i + \beta_5 gen_i \\ + \beta_6 sta_i + \beta_7 inc_i + \beta_8 book_i + \beta_9 hour_i + r_{ij} \qquad (7)$$

Level 2:
$$\beta_{0j} = \gamma_{00} + \gamma_{01} type + \gamma_{02} moda_j + \gamma_{03} reg_j + u_{0j} \qquad (8)$$

where: $Enade_{ij}$ is the score of student i from the undergraduate program in Accounting j on the 2015 Enade; β_{0j} is mean Enade score for all students in program j; $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9$ are the level-1 regression parameters; *CN, CH, LT* and *MT* are continuous variables for previous achievement (Enem scores); *gen* is the dummy variable for gender; *sta* is the dummy variable for marital status; *inc* is the dummy variable for income; *book* is the dummy variable for number of books students read a year; *hour* is the dummy variable for hours used for extra-class studies; r_{ij} is the residual score of student i in program j; γ_{00} is the overall mean Enade score of all undergraduate programs in Accounting in Brazil; γ_{01} is the difference in the mean Enade scores between programs in universities/university centers and programs in colleges; γ_{02} is the difference in the mean Enade scores between the in-person learning programs and distance learning programs; γ_{03} is the difference in the mean Enade scores between programs in the South and Southeast and programs in the other regions of Brazil; u_{0j} is the residual score of program j, pointing to a difference between program j and the overall mean Enade score.

Regression analysis was performed twice. The first regression served to estimate the parameters, observe the predicted values and estimate the standardized residuals. The second regression served to exclude outliers, i.e., those values with standardized residuals above | 3 |. The data reported in this research were obtained in the model after excluding outliers. The parameters values withdrawn were used to estimate $\widehat{Enade}_{ij}$:

$$\widehat{Enade}_{ij} = \hat{\beta}_{0j} + \hat{\beta}_1 CN_i + \hat{\beta}_2 CH_i + \hat{\beta}_3 LT_i + \hat{\beta}_4 MT_i + \hat{\beta}_5 gen_i + \hat{\beta}_6 sta_i \qquad (9)$$
$$+ \hat{\beta}_7 inc_i + \hat{\beta}_8 book_i + \hat{\beta}_9 hour_i$$

The gross IDD-VDCF$_{ij}$ is calculated for each student *i* of the undergraduate course *j* (Equation 6); then the average per course is calculated. This value is standardized and then staggered, ranging from 1 to 5, following the same guidelines as Technical Note No. 17/2018 that regulates the calculation of IDD-Inep (INEP, 2017).

The statistical procedures used for the IDD-VDCF Model were: (1) multicollinearity analysis of the explanatory variables; (2) backward method for extracting significant variables; (3) Likelihood Ratio Test to choose the best model (TRV); (4) testing assumptions about residues; (5) calculation of added value. The parameters of the IDD-VDCF model were estimated using the likelihood method.

Deviance statistics, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were used to compare the proposed model (IDD-VDCF) with the current model of Inep. The Deviance statistic measures the model's degree of mismatch, so

that when comparing models, the lower the Deviance value, the better the adjustment obtained (LAROS; MARCIANO, 2008). The Deviance calculation is given by D = -2log (L), where L is the model's likelihood value. The Akaike Information Criterion (AIC) is given by AIC = -2 log (L) + 2p, where p is the number of model parameters (AKAIKE, 1974). The Bayesian Information Criterion (BIC) is given by BIC = -2 log (L) + 2plog (n), where n is the total number of observations (SCHWARZ, 1978).

5 Results

5.1 The adequacy of proposed model

Following Laros and Marciano (2008), the first decision was to test whether the multilevel model (HLM) would be the most appropriate to assess value-added in Brazilian undergraduate programs in Accounting. To this end, a null model was tested, i.e., without explanatory variables (see equations 10 and 11).

$$\text{Nível 1:} Enade_{ij} = \beta_{0j} + r_{ij} \tag{10}$$

$$\text{Nível 2:} \beta_{0j} = \gamma_{00} + u_{0j} \tag{11}$$

The results of the null model supported the use of the multilevel model, since the intraclass correlation (ICC) was 0.17, i.e., 17% of the total variance in the Enade scores is explained by differences between the programs. The second stage was to assess the multicollinearity between the explanatory variables to ensure the model is accurate in only estimating parameters that are not highly correlated. This included assessing the Variance Inflation Factor (VIF): the higher the coefficient of determination, the higher the VIF indicating high collinearity between variables. No variables had VIF above 10, meaning no one should be removed a priori (HAIR *et al*., 2010).

The first step was to include the level-1 (student) variables. Table 1 summarizes the results of the fixed effects model with random intercept for the programs. At the student level, prior knowledge measured through Enem scores and all sociodemographic variables tested in the model, were significant predictors of Enade achievement.

Table 1 - IDD-VDCF model estimates with level-1 variables

	Estimate	Standard Error	Degrees of Freedom	t-value	p-value
Fixed Effects (Level 1)					
Intercept	-4.17	0.60	29630	-7.01	0.00
Enem CN	0.02	0.00	29630	18.94	0.00
Enem CH	0.02	0.00	29630	15.82	0.00
Enem LT	0.04	0.00	29630	29.21	0.00
Enem MT	0.02	0.00	29630	32.28	0.00
Gender	-1.64	0.12	29630	-13.61	0.00
Marital status	-0.43	0.13	29630	-3.34	0.00
Income	-1.09	0.13	29630	-8.50	0.00
Books	-0.28	0.12	29630	-2.40	0.02
Hours	-1.25	0.12	29630	-10.55	0.00

Source: research data.

The coefficient of determination (R^2), which measures the fit of the model, was 37.90%. The ICC was 10.09%, which resulted from including the Enem-related variables. Liu (2011b) says that the ICC reduced from 15 to 10% on average when prior knowledge is controlled in the model. The Bayesian Information Criterion (BIC) was 226621, which is lower than the value in the null model, i.e., indicates a better fitted model.

The findings consistently show that the percentage of variance explained by the programs was reduced (ICC = 0.09) after introducing student and program level explanatory variables into the equations. The main reason is that the Enem scores accounted for a substantial amount of variation in the Enade achievement because of the high correlation between both Enade and Enem scores ($\rho = 0.55$). The adjusted determination coefficient of the final model was 38.05% as shown in Table 2.

Table 2 - Statistics of final Model IDD-VDCF

Fixed Effect		Estimate	Standard Deviation	t-value	p-value
Intercept		-9.82	1.08	-9.06	0.00
Enem CN		0.02	0.00	18.79	0.00
Enem CH		0.02	0.00	15.80	0.00
Enem LT		0.03	0.00	29.05	0.00
Enem MT		0.02	0.00	31.98	0.00
Gender		-1.67	0.12	-13.84	0.00
Marital status		-0.48	0.13	-3.71	0.00
Income		-1.05	0.13	-8.14	0.00
Number of books		-0.29	0.12	-2.48	0.01
Hours of study		-1.26	0.12	-10.58	0.00
Academic organization		1.01	0.26	3.82	0.00
Learning modality		5.02	0.92	5.48	0.00
Region of the undergraduate program		0.92	0.27	3.42	0.00
Random effects					
Intercept variance		9.52	3.08		
Residual variance		93.67	9.68		
Intraclass correlation		0.09			
Number of observations		30668			
Number of groups		911			
Goodness of fit	AIC	226475		Adjusted R^2	38.05
	BIC	226599			

Source: research data.

Table 2 shows that students who read up to three books a year have on average achievements 0.29 lower on the Enade than those who read above three books. Students who study up to three hours a week have on average 1.26 lower achievements on the Enade than those studied for three or more hours. Women perform lower than men (1.67 on average), and singles also score lower (0.48 on average). Students with family income of up to 3,258.00 BRL have an Enade score on average 1.05 lower than those with incomes that equaled or exceeded that amount.

Regarding institutional characteristics, students enrolled in programs in universities, university centers or federal institutes have mean Enade scores 1.01 higher than those peers in colleges. In-person learning programs students have Enade scores on average 5.02 higher than those students in distance learning programs. Students attending programs in the South and Southeast of the country have mean Enade scores 0.92 higher than the others.

To assess reliability, residual analysis was performed through graphs and statistical tests. The residuals have a normal distribution, which was confirmed by a Kolmogorov-Smirnov normality test with p-value of 0.08 (i.e., above the significance level of 0.05). In other words, the statistical test does not reject the null hypothesis that the residuals have

normal distribution. The F test was also performed and confirmed homogeneity in the variance of the model residuals (p-value = 0.44).

After identifying the best fit regression equation for the context of Accounting programs in Brazil, the IDD in Model IDD-VDCF was analyzed by estimating the difference between the students' actual achievements and expected achievements, considering their admission characteristics (Enem scores), their personal characteristics (gender, marital status, income, reading and study habits) and the conditions of learning (type of academic organization, type of education, and regional location of the program).

5.2 Comparison between IDD-VDCF Model and Inep Model

Model IDD-VDCF is different from the Inep model in that it includes explanatory variables for achievement at both level 1 (student) and level 2 (programs). At level 1, the terms $\beta_5(gen_i)$, $\beta_6(sta_i)$, $\beta_7(inco_i)$, $\beta_8(book_i)$ and $\beta_9(hour_i)$ were introduced in the model as predictors of achievement. At level 2, the terms $\gamma_{01}(type_j)$, $\gamma_{02}(moda_j)$ and $\gamma_{03}(reg_j)$ were also included in the equation. Fernandes, Miranda and Alexander (2020) have shown that Inep model could be improved by introducing explanatory variables which are not controlled by higher education institutions.

The correlation between the models compared is 0.94, which shows a high correspondence between the IDDs in both models. Both models have significantly different means ($\alpha=5\%$), even though their correlation is high (see Table 3).

Table 3 - T-test for dependent samples (Model Inep and Model IDD-VDCF)

	Inep	IDD_VDCF
Mean	2.36	2.54
Variance	0.21	0.20
Pearson Correlation	0.94	
t-statistics	-35.94	
p-value	0.00	

Source: research data.

Statistically, Model IDD-VDCF provides a better fit of the data compared to the Inep model, as it presents lower values of AIC and BIC, as shown in the analysis of variance between both models (see Table 4). In addition, the Likelihood Ratio Test (LRT) was performed, in what:

H0: the simplest model (Inep) fits as well as the Model IDD-VDCF

H1: Model IDD-VDCF fits significantly better than Model Inep.

Table 4 - Variance Analysis Between the Inep and IDD-VDCF Models

Model	DF	AIC	BIC	logLik	p-value
IDD-VDCF	15	228647.4	228772.3	-114309	
Inep	7	229096.5	229154.8	-114541	<.0001

Source: research data.

Through the LRT ($\alpha=5\%$), the null hypothesis is rejected, that is, the IDD-VDCF Model presented the best fit. Furthermore, the coefficient of determination of the proposed model ($R2 = 38.05$) is higher than the coefficient of Inep Model ($R2 = 37.05$). Another indication that, statistically, the proposed model is more adequate to the analyzed database is the mean square of the error (QME), in which in the IDD-VDCF Model it was 91.9, while in the Inep model it was 97.6.

The proposal of a new value-added estimation model (IDD-VDCF) for undergraduate programs in Accounting in Brazil follows the identification of the determinants of academic achievement. Bailey and Xu (2012) contend that factors unrelated to achievement should be investigated if one is to provide comprehensive, unmistaken data on institutional effectiveness, especially when those factors are beyond institutional control. As students are admitted to undergraduate programs with varying degrees of academic skills and aspirations, these personal characteristics can affect their achievement, including their likelihood of completing their studies (BAILEY; XU, 2012).

Liu (2011a, 2011b) argues that it is also necessary to control institutional variables, such as the selection process and the provision of graduate programs, which may affect student's final achievement. Steedle (2012) refutes this claim: in his study, the control variables (administrative category, historically black people-oriented HEIs, existence of graduate programs, selectivity, full-time enrolment, percent enrolled white, full-time retention rate, overall graduation rate, and student/faculty ratio) accounted for 10% of the variance in the achievement scores, but none of them was significant in the model after introducing the previous student achievement at the institution level.

5.3 Practical Implications of IDD-VDCF Model

Assuming that the purpose of programs, faculty, organization leaders and policy makers is to provide quality education that adds to the students' academic development, greater importance should be placed on analyzing a program's value-added alongside student achievement. In other words, it is necessary to find out which variables are related to a

program's value-added (IDD), so that the program coordinators can provide better schooling to their students.

To this end, a correlation test was performed between the student and program related variables and the different IDD estimation models (IDD-VDCF Model and Inep Model), assuming that an understanding of the determinants of academic gain will support organization leaders in making practical decisions to improve the institutional effectiveness of their programs. In Table 5, the IDD is significantly correlated with individual student characteristics in Inep model. As a program's value-added is correlated with the students' sociodemographic characteristics regardless of institutional efforts, it follows that the public policies of social inclusion are essential in Brazil. If the IDD conceptually is an indicator that measures the value a program adds to student achievement (INEP, 2017), it should not be related to the students' personal characteristics, as in IDD-VDCF. The proposed model has only one significant correlation with individual characteristic (Enem MT), suggesting that it does a good job of capturing programmatic differences not individual characteristics. Controlling for academic characteristics is crucial if the goal is to measure a program's contribution. Not surprisingly, the program-related variables, such as the characteristics of faculty members' (doctoral degree, master's degree, and employment contract), are significantly related to the programs' contribution to their students' academic achievement. Since these are mutable factors, these correlations are fair and can offer insight on how to make programmatic improvements.

In Table 5, the correlation coefficients in Inep Model are higher than those in IDD-VDCF Model because including control variables reduced the effect of institutional characteristics on academic achievement. This is another advantage of IDD-VDCF over Inep Model; after all, policy makers want to identify which institutions are producing higher performers but reporting only differences in scores may be insufficient because of the significance of program-related variables.

Table 5 - Correlation between IDD and student characteristics

Student Characteristics		IDD_Inep	IDD_VDCF	Programs characteristics		IDD_Inep	IDD_VDCF
Marital status	ρ_{pb}	-0.01	0.04*	Academic organization	ρ_{pb}	0.11*	0.03
Ethnicity	ρ_{pb}	0.01	0.04*	No. of programs in the HEI	ρ	0.24*	0.18*
Nationality	ρ_{pb}	0.00	0.01*				
Level of education (father's)	ρ_{pb}	-0.00	0.00	Learning modality	ρ_{pb}	0.05	-0.03
Level of education (mother's)	ρ_{pb}	0.00	0.00	(ODP) Pedagogical Teaching Structure	ρ	0.26	0.25
Income	ρ_{pb}	0.01	0.03*	(IFF) Infrastructure	ρ	0.31*	0.29*
Work	ρ_{pb}	0.01	0.02*	(OAF) Opportunity for further learning	ρ	0.30*	0.27*
Funding	ρ_{pb}	0.00	0.01	Faculty with a master's degree (Me)	ρ	0.18*	0.16*
Affirmative policy	ρ_{pb}	0.00	-0.00	Faculty with a doctoral degree (Doc)	ρ	0.11*	0.10*
High school institution	ρ_{pb}	-0.01	0.01	Employment contract (RT)	ρ	0.16*	0.16*
Type of high school	ρ_{pb}	0.01	0.02*				
First in the family in an HEI	ρ_{pb}	-0.01	-0.01				
No. of books read	ρ_{pb}	.000	-0.02*				
Hours of study	ρ_{pb}	-0.00	-0.05*				
Enem (CN)	ρ	0.01	0.04*				
Enem (CH)	ρ	0.00	0.03*				
Enem (LT)	ρ	0.00	0.04*				
Enem (MT)	ρ	0.01*	0.06*				

Source: research data.

Program coordinators have little control over such sociodemographic variables, since public policies for social inclusion are in general at the level of the HEIs or the federal government. Therefore, attention should be drawn to the fact that the number of books read and hours of study are related to academic achievement. As programs with students who read more books and study longer outperforms the others, practical measures that encourage such activities can improve institutional effectiveness. The present data also show that pedagogical teaching structure, opportunity for further learning and program infrastructure are variables significantly related to a program's value-added, i.e., the higher the rates in those variables, the greater a program's contribution to student achievement on the Enade.

6 Discussion

This article discusses important issues for education: the methodology for measuring achievement in higher education and the possibility/feasibility of equally estimating the achievement among students with very different socioeconomic backgrounds (a broad discussion that is far from over). The topic value-added is particularly relevant in Brazil, and also in other countries where the wide variation in income and the level of access to knowledge at the most basic levels of education are social issues. Similar discussions have

been taking place on the axis North (in the USA and Europe, in particular), it has been found lately that these socioeconomic discrepancies are also realities to consider in these regions, which highlights the importance of this discussion also in these contexts.

An important operational concern in value-added models is identifying what quality differences among students to consider in order not to underestimate or overestimate the value-added. Shavelson *et al*. (2016) argue that students from better socioeconomic backgrounds can travel, speak other languages, and more easily find internship opportunities due to their parents' networking. Ignoring such characteristics could disadvantage institutions that admit students from low socioeconomic backgrounds when compared to those with more privileged students.

On this issue, the pandemic in 2020/2021 opens the socio-economic gap between students due the challenges faced by all universities across the world: the shifting from face-to-face to online classes, new methods to assessment and evaluation, travel restrictions, mental health, student support services to deal with the crisis, among others. If, on the one hand, the closing of schools contributes to the reduction of the number of deaths by Covid-19, on the other, adverse effects are evidenced mainly among more vulnerable families: low quality internet access , insufficient number of devices to access classes at home (FRENETTE; FRANK; DENG, 2020); depression and anxiety (RUDENSTINE *et al*., 2021); school dropout, job loss and increased school debt (ONYEMA *et al.,* 2020); increase in inequality in standardized tests (HAECK; LEFEBVRE, 2020) Therefore, privileged students have more access to adequate distance learning conditions, so that, in the future, the tendency may be for this gap to widen further. Therefore, considering the SES seems an inexorable condition to discuss value-added.

The VAM proposed for Brazilian context admits that there are differences at the moment of admission, but also in the conditions of learning; therefore, it is necessary to level the students based on their personal characteristics in order to isolate this effect from the school contribution. Although IDD-VDCF Model is statistically close to the Inep model, it seeks to reduce discrepancies between programs by controlling for sociodemographic and institutional variables. The IDD-VDCF distinguishes itself for its lack of significant associations with any of the individual characteristics in contrast to the Inep Model. The residual related to each program, which allows for ranking the institutions, may carry the weight of unobservable variables, but admittedly capable of influencing academic achievement, such as student-teacher relationship, dedication of program coordinator, learning environment, use of different methodologies in the classroom. These variables are

partially captured by the variables measured in the Sinaes: infrastructure, pedagogical teaching structure, and opportunity for further learning.

The findings draw attention to the social structure of undergraduate programs in Brazil. Even though such findings are representative only of the undergraduate programs in Accounting, they may be similar for programs in several other fields. Still, a model was developed that controls for the students' socioeconomic conditions, this research reveals that inequality persists in HE. Despite the public policies of social inclusion headed by the Worker's Party government between 2003 and 2018, the most privileged students have still dominated the best programs in the country. As such, the higher a program's ranking, the better its institutional effectiveness (i.e., the higher the value-added to student achievement). Yeh (2020) would consider this the demoralizing effect while Merton (1968, p. 3) refers to this as the Mathew effect: "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath", i.e. those with more get more.

Yeh (2010), when studying K-12 education, identified that small differences between students' previous achievement are amplified by the structure of schools, not only perpetuating, but widening the achievement gap that exists between poor students and their wealthier colleagues. In this sense, Yeh (2020) argues that the application of standardized tests, classification and comparison of groups based on them serves to demoralize underperforming students, thus justifying the permanence of the achievement gap between low-income minority students and their more affluent peers.

The CPC is the major indicator used by the federal government to allocate public resources in Brazil and its results do matter to organization leaders. Public funding of HEIs has not been based on the institutions' effectiveness or quality, but rather on their size, which generally relates to the magnitude of personnel spending (CANZIANI *et al*., 2018). According to Johnes (2018) producing a single achievement measure from various dimensions results in complexity in interpreting this measure since appropriate weights are necessary to combine the previously separate measures. In general, the weights are arbitrary for institution ranking, causing discrepancies for a given institution or program across different systems (USHER; MEDOW, 2009), as reported by Johnes (2018) for the different U.S. ranking systems. For this reason, combining several indicators into a single measure may be questionable. In most countries, where no national assessment system is in place, different indicators attest to the quality of an undergraduate program, which comes back to the question of values and purpose of education discussed in Mitchell and Mitchell (2003).

Therefore, reanalyzing the weights of the variables that make up the CPC is necessary both for higher education in general and for undergraduate programs in Accounting in particular. Building on the data analyzed in this research, it seems that the quality indicator CPC is in fact reinforcing the already existing difference between the programs as it uses the IDD and the conditions of the learning to estimate a single measurement of quality. As shown in this research the value-added is positively related to these characteristics. In other words, the higher the IDD, the higher the Program Quality Levels for faculty and infrastructure.

Acknowledgements

References

AKAIKE, H. A new look at the statistical model identification, **IEEE Transactions on Automatic Control**, New Jersey, v. 19, n. 6, p. 716-723, 1974. Available at: https://ieeexplore.ieee.org/document/1100705. Accessed on: 3 Aug. 2019.

ALEXANDER, Nicola; JANG, Sung Tae; KANKANE, Shipi. The performance cycle: the association between student achievement and state policies tying together teacher performance, student achievement, and accountability. **American Journal of Education**, Chicago, v. 123, n. 3, p. 413-446, 2017. Available at: https://doi.org/10.1086/691229. Accessed on: 9 Oct. 2019.

ALEXANDER, Nicola; JANG, Sung Tae. Expenditures on the professional development of teachers: The case of Minnesota. **Journal of Education Finance**, Baltimore, v. 44, n. 4, p. 385-404, 2019. Available at: https://www.muse.jhu.edu/article/738161. Accessed on: 9 Oct. 2019.

BAILEY, Thomas; XU, Di. Input-adjusted graduation rates and college accountability: what is known from twenty years of research? **Context for Success**, Austin, 2012. Available at: http://www.hcmstrategists.com/contextforsuccess/papers/LIT_REVIEW.pdf. Accessed on: 10 Sep. 2020.

BOGOYA, José Daniel; BOGOYA, Johan Manuel. An academic value-added mathematical model for higher education in Colombia. **Ingeniería E Investigación**, Bogotá, v. 33, n. 2, p. 76-81, 2013. Available at: https://revistas.unal.edu.co/index.php/ingeinv/article/view/39521/42363. Accessed on: 19 Aug. 2019.

BRASIL. **Lei nº 10.861, de 14 de abril de 2004.** Institui o Sistema Nacional de Avaliação da Educação Superior – SINAES e dá outras providências. Brasília, 2004. Available at:

http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.861.htm. Accessed on: 10 Sep. 2020.

CANZIANI, Alex *et al.* (org.). **Financiamento da educação superior no Brasil**: impasses e perspectivas. Brasília: Câmara dos Deputados, 2018. Available at: https://www2.camara.leg.br/a-camara/estruturaadm/altosestudos/pdf/financiamento-da-educacao-superior-no-brasil-impasses-e-perspectivas. Accessed on: 10 Sep. 2020.

CUNHA, Jesse; MILLER, Trey. Measuring value-added in higher education: possibilities and limitations in the use of administrative data. **Economics of Education Review**, Amsterdam, v. 42, p. 64-77, 2014. Available at: https://doi.org/10.1016/j.econedurev.2014.06.001. Accessed on: 10 Sep. 2019.

FERNANDES, V. D. C. *et al*. ENADE: Uma análise sobre os determinantes do desempenho acadêmico dos estudantes de Ciências Contábeis desde a sua primeira edição. *In*: ENCONTRO NACIONAL DA ANPAD, 42, 2018, Curitiba. **Anais** [...]. Curitiba: ANPED, 2018. p. 1-16.

FERNANDES, V. D. C.; MIRANDA, G. J.; ALEXANDER, N. Value-added measures in higher education: a historical contextualization of Brazilian experiences. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 101, n. 259, p. 691-720, 2020. Available at: https://doi.org/10.24109/2176-6681.rbep.101i259.4469. Accessed on: 10 Feb. 2021.

FRENETTE, M.; FRANK, Kristyn; DENG, Zechuan. School closures and the online preparedness of children during COVID-19 Pandemic. **Economics Insight: Statistics Canada**, Ottawa, n. 103, 2020. Available at: https://files.eric.ed.gov/fulltext/ED605398.pdf Accessed on: 18 Jun. 2020.

HAECK, C.; LEFEBVRE, P. Pandemic school closures may increase inequality in test scores. **Canadian Public Policy**, Toronto, v. 26, n.1, p. 82-87, jul. 2020. Available at: https://doi.org/10.3138/cpp.2020-055. Accessed on: 18 Nov. 2020.

HAIR, Joseph *et al*. **Multivariate data analysis**. 7. ed. New Jersey: Prentice Hall, 2010.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **NOTA TÉCNICA Nº 33/2017/CGCQES/DAES.** Brasília: Inep, 2017. Available at: http://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2016/nota_tecnica_n33_2017_cgcqes_daes_calculo_idd.pdf. Accessed on: 10 Aug. 2020.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **ENADE**. Brasília: Inep, 2018. Available at: http://portal.inep.gov.br/web/guest/enade. Accessed on: 10 ago. 2020.

INEP - INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Censo da educação superior:** notas estatísticas 2018. Brasília: Inep, 2019. Available at: http://download.inep.gov.br/educacao_superior/censo_superior/documentos/2019/censo_da_educacao_superior_2018-notas_estatisticas.pdf. Accessed on: 10 ago. 2020.

JOHNES, Jill. University rankings: what do they really show? **Scientometrics**, Budapest, v. 115, n. 1, p. 585-606, 2018. Available at: https://doi.org/10.1007/s11192-018-2666-1. Accessed on: 19 Sep. 2019.

KIM, Hoonho; LALANCETTE, Diane. **Literature review on the value-added measurement in higher education**. Paris: OECD. 2013. Available at: https://www.oecd.org/education/skills-beyond-school/Litterature%20Review%20VAM.pdf. Accessed on: 19 Apr. 2020.

LAROS, Jacob Arie; MARCIANO, João Luiza Pereira. Análise multinível aplicada aos dados do NELS:88. **Estudos em Avaliação Educacional**, São Paulo, v. 19, n. 40, p. 263-278, 2008. Available at: https://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1440/1440.pdf Accessed on: 24 Nov. 2019.

LIU, Lydia. Measuring value-added in higher education: conditions and caveats - results from using the Measure of Academic Proficiency and Progress (MAPP^{TM}). **Assessment and Evaluation in Higher Education**, London, v.36, n. 1, p. 81-94, 2011a. Available at: https://doi.org/10.1080/02602930903197917. Accessed on: 12 May 2019.

LIU, Lydia. Value-added assessment in higher education: a comparison of two methods. **Higher Education**, Berlim, v. 61, p. 445-461, 2011b. Available at: http://dx.doi.org/10.1007/s10734-010-9340-8. Accessed on: 12 May 2019.

MELGUIZO, Tatiana *et al*. The methodological challenges of measuring student learning, degree attainment, and early labor market outcomes in higher education. **Journal of Research on Educational Effectiveness**, London, v. 27, n.2, p. 424-448, 2017. Available at: https://doi.org/10.1080/19345747.2016.1238985. Accessed on: 18 Jul. 2019.

MERTON, R. K. The Matthew Effect in Science: the reward and communication systems of science are considered. **Science**, Washington, v. 159, n. 3810, p. 56-63, 1968. Available at: https://doi.org/10.1126/science.159.3810.56. Accessed on: 12 Nov. 2019.

MILLA, Joniada; SAN MARTÍN, Ernesto; BELLEGEM, Sébastien. Higher education value added using multiple outcomes. **Journal of Educational Measurement**, Boston, v. 53, n.3, p. 368-400, 2016. Available at: https://doi.org/10.1111/jedm.12114. Accessed on: 11 Jun. 2019.

MITCHELL, Douglas; MITCHELL, Ross. The political economy of education policy: the case of class size reduction. **Peabody Journal of Education**, London, v. 78, n.4, p. 120-152, 2003. Available at: https://doi.org/10.1207/S15327930PJE7804_07. Accessed on: 2 Aug. 2019.

NATIONAL RESEARCH COUNCIL. **Getting Value out of Value-Added:** report of a Workshop. Washington, DC: The National Academies Press, 2010.

ONYEMA, E. M. *et al.* Impact of Coronavirus Pandemic on education. **Journal of Education and Practice**, USA, v. 11, n. 13, p. 108-121, 2020. Available at: https://iiste.org/Journals/index.php/JEP/article/view/52821/54575. Accessed on: 18 Dec. 2020.

PIKE, G. Considerations when using value-added models in higher education assessment. **Assessment Update**, London, v. 28, n. 5, p. 8-10, 2016. Available at: https://doi.org/10.1002/au.30073. Accessed on: 2 Sep. 2019.

RUDENSTINE, S. *et al*. Depression and Anxiety During the COVID-19 Pandemic in an Urban, Low-Income Public University Sample. **Journal of Traumatic Stress**, Boston, v. 34, p. 12-22, 2021. Available at: https://doi.org/10.1002/jts.22600. Accessed on: 10 Jan. 2021.

SCHWARZ, G. Estimating the Dimension of a Model. **Ann. Statist.**, Waite Hill, v. 6, n. 2, p. 461-464, 1978. Available at: https://doi.org/10.1214/aos/1176344136. Accessed on: 3 Aug. 2019.

SHAVELSON, Richard *et al*. On the practices and challenges of measuring higher education value-added: the case of Colombia. **Assessment and Evaluation in Higher Education**, London, v. 41, n. 5, p. 695-720, 2016. Available at: https://doi.org/10.1080/02602938.2016.1168772. Accessed on: 5 Jul. 2019.

STEEDLE, J. T. Selecting value-added models for post-secondary institutional assessment. **Assessment and Evaluation in Higher Education**, London, v. 37, n. 6, p. 637-652, 2012. Available at: https://doi.org/10.1080/02602938.2011.560720. Accessed on: 5 Jul. 2019.

USHER, Alex; MEDOW, Jon. A global survey of university rankings and league tables. *In*: KEHM, B. M.; STENSAKER, B. **University rankings, diversity, and the new landscape of higher education**. Rotterdam: Sense Publishers, 2009. p. 3-18.

YEH, Stuart. Understanding and addressing the achievement gap through individualized instruction and formative assessment. **Assessment in Education: Principles, Policy and Practice**, London, v.17, n.2, p. 169-182, 2010. Available at: https://doi.org/10.1080/09695941003694466. Accessed on: 5 Jul. 2019.

YEH, Stuart. Educational accountability, value-added modeling, and the origin of the achievement gap. **Education and Urban Society**, Los Angeles, v. 52, n. 8, p. 1181-1203, 2020. Available at: https://doi.org/10.1177/0013124519896823. Accessed on: 10 Feb. 2020.