ISSN: 2230-9926

Available online at http://www.journalijdr.com

IJDR

International Journal of Development Research
Vol. 12, Issue, 03, pp. 54795-54798, March, 2022
https://doi.org/10.37118/ijdr.24160.03.2022

RESEARCH ARTICLE OPEN ACCESS

THE CURRENT ASSESSMENT OF THE STUDENT'S ACADEMIC ACHIEVEMENT IS A BIG MISTAKE

Cristiano Mauro Assis Gomes[1,*], Enio Galinkin Jelihovschi[2] and Jhonys de Araujo[1]

[1]Laboratory for Cognitive Architecture Mapping (LAICO), Universidade Federal de Minas Gerais (UFMG), Department of Psychology, Campus Pampulha, Av. Pres. Antônio Carlos, 6627 – Pampulha, Belo Horizonte – MG, Brazil, CEP: 31270-901

[2]Universidade Estadual de Santa Cruz (UESC), Department of Exact and Technological Sciences, Campus Soane Nazaré de Andrade, Nazaré de Andrade Rodovia Ilhéus-Itabuna km. Ilhéus – BA, Brazil, CEP: 16.45.662-900

ARTICLE INFO

Article History:
Received 05th January, 2022
Received in revised form
28th January, 2022
Accepted 17th February, 2022
Published online 28th March, 2022

Key Words:
Student Assessment,
Measurement, Achievement.

****Corresponding author:***
Cristiano Mauro Assis Gomes

ABSTRACT

This study argues that the current assessment of the student's academic is wrong. The educators, in general, were trained in a secular practice that the student' academic achievement can be evaluated through the measurement of the population or a group of students. This direct transposition of the estimate of the population to one particular student is the reason why the current assessment is wrong. This study shows the main arguments that sustain this statement, cautioning educators that a sweeping transformation is necessary for a proper evaluation of the academic achievement and learning of the student. In sum, in this article we show that all we know about the student's achievement and learning, including their predictors, are not valid. Furthermore, we also claim that researchers and educators should recognize the importance of this issue and consider other alternatives to render valid and viable the measurement of the student's achievement and learning.

Citation: ***Cristiano Mauro Assis Gomes1, EnioGalinkin Jelihovschi2 and Jhonys de Araujo1***. "The Current Assessment of the Student's Academic Achievement is a Big Mistake*", International Journal of Development Research*, 12, (03), 54795-54798.

INTRODUCTION

The educational field has been guided by a secular postulate that it is possible to assess and predict the student's academic achievement using directly the estimations based on population or groups of students. This postulate is so old and well-established that it is almost perceived as an unquestionable truth and gives support to a secular practice of the student's academic achievement inference based on the estimate of the academic achievement of a population or group of students. This practice occurs widely, from the current evaluations in the schools to the large-scale educational tests, which are, at least in theory, sophisticated and well-designed tools with robust evidence of validity. For example, in Brazil, a large-scale test which evaluates the students at the end of secondary education is the National Exam of Upper Secondary Education (*Exame Nacional do Ensino Médio* [Enem]).

Enemis a very important educational exam in Brazil because its scores are used as the gateway to the most prestigious higher education institutions in this country. Enemis applied once a year and millions of students perform the exam. The Enemgenerates four scores for each student, math score, language score, natural science score, and human science score (Gomes, Golino *et al*., 2018). Each student has a score in these four domains and this information is used as an inference about the knowledge acquired by the student at the end of the secondary education. The Enem's scores are produced through an item response theory model that estimates its parameters based on the population that performs the exam. In fact, from the statistical point of view, the score of each individual is considered mere replication of the same random variable. This is the same as taking as a hypothesis that individuals are homogeneous in all relevant respects of analysis.

Therefore, the student does not have her (his) own estimate. This is a very ubiquitous point of in all science of psychological analysis and testing.

The Apparent Truth is a Big Mistake: The Molenaar's (2004) manifesto shows that the social and human sciences make a big mistake when they estimate their theoretical constructs using populations or groups of individuals and direct apply this estimation to the individual. Molenaar (2004) states that someone only can transpose the information based on population to the individual if this population is ergodic. All processes followed by a student learning a new subject, learning a test or even a child learning how to write or any other kind of psychological process, can be modeled mathematically by what is called a stochastic process. The importance of a mathematical model is that it permits the researcher to figure out how the results of the events, as the process goes along, are going to behave in most of the cases. The word stochastic means randomness since the results are random events, that is, there is always a chance that an event will happen in any other way. The theory of stochastic processes, as part of the whole probability theory, have been studied for more than a hundred years and has become the best tool to study and to interpr*et al*l kinds of psychological processes.

Let us now go back to the problem we are dealing with in this paper. Many students are going through the same learning process and the amount of whatever they learn is being repeatedly measured in time. The result of this measure is called a state of the process, and the performance is defined by the average learned against an average measurement which has been established using results from the whole population of students. A fundamental question arises in this practice: can this population average measurement be applied to every student? In other words, how can we be sure that each student's results, taken as a random variable, has the same theoretical average measurement as the population average? Why is this problem so fundamental? Because if we are not sure, we cannot calculate the performance as we did above, since, in theory, each student average is different from the population average and therefore they cannot be compared. Let us explain it in other words. If in theory the student average were equal to the population average, the sample average resulting from this student would drift around that average with no bias. On the other hand, if in theory the students average differs from the population average the comparison will be very biased, that is, the former will be constantly below or above the later and the calculated performance will have no meaning. We can only solve this fundamental question by going back to a mathematical model and to the theory of stochastic processes.

First of all, we need to describe a special type of stochastic process called ergodic. Those ergodic processes use three concepts. Stationarity, absorbing state and cycle. A stationary process has the property that the mean, variance and autocorrelation structure do not change over time, in other words all the probability laws that govern the process do not change over time. This means that, in a long run, the process forgets its initial state, which could be more erratic, and becomes homogeneous in time. In this context, it does not matter anymore where and how it started. Absorbing state is the name of a state where the process, after entering it can no longer exit and go to another state. For instance, suppose a drunk man has wandered far too close to a cliff. From where he stands, one step forward would send the drunk man over the edge. He takes random steps, either towards or away from the cliff. At any step, the probability of taking a step away is 2/3 and a step towards the cliff is 1/3. This is a stochastic process and the possible states are the positions he can step in with respect to the cliff. From every position he has a probability to reach another one, except one position, the cliff. If he falls from the cliff he would no longer get back. The cliff is an absorbing state of this process. A cycle is a collection of states where the process circulates around without escaping. Once inside the cycle, it is not possible to reach any other state besides those of the cycle. An ergodic process is a stationary process which has neither an absorbing state nor a cycle. The so-called Brownian motion is an ergodic process. Brownian motion is the movement that a small particle, like a pollen, performs on the surface of water as the moving molecules of the liquid bump against it. If we describe the process as the graphical movement of the particle, it can be described as an ergodic process. It is homogeneous, stationary, it has no absorbing state since it keeps moving around the whole surface and also has no cycle, for the same reason. The following theorem, called the Birkhoff (1931) Ergodic Theorem solves our fundamental question. Heuristically, the theorem expresses the conditions under which a process will have, in a long run, an average along the path of one element evolving in time through the states of the process to be equal to the average of all possible states which the process may attain at any given time. That is a process which are very homogeneous along the time as it is along the space. The sole condition is: the process must be ergodic. Molenaar (2004), in that seminal manifesto realized the importance of the ergodic theorem and claimed that psychology as an idiographic science has to bring the person back to the center of psychological research. Moreover, in any psychological testing the IEV-interindividual variation which is the variation between individuals and IAV-intraindividual variation which is the variation in a single participant's time series can never yield the same average given the extreme variability in human genetics and cultural upbringing. This result shows that psychological testing can never be ergodic.

Considering that ergodicity is very difficult to occur in academic achievement, then it is a big mistake to directly transpose information based on population to the individual when trying to infer the achievement of a student. In consequence, it is mandatory to estimate the own student's achievement to infer about it. Until now, the evidence inform us that the structure found in the individual is different from the evidence found in the population (Gomes, Araujo, *et al*., 2018). The direct transposition of the estimate based on population to the student is very common. This occurs in the evaluation, as well as, in the prediction of the student's academic achievement. The evidence about the predictors of the academic achievement are plentifully based on estimate of the population, such as intelligence (Gomes, 2010b, 2011b, 2012b; Gomes&Borges, 2007, 2008b, 2009b, 2009c), personality (Gomes, 2012a; Gomes &Gjikuria, 2017), socioeconomics variables (Gomes &Almeida, 2017; Gomes, Amantes, *et al*., 2020; Gomes &Jelihovschi, 2019; Gomes, Lemos *et al*., 2020), students' approaches to learning (Gomes, 2010c, 2011a, 2013; Gomes, Araujo, *et al*., 2020), students' beliefs on teaching-learning processes (Gomes &Borges, 2008a), motivation for learning (Gomes &Gjikuria, 2018), self-regulatory abilities, and mediated processes of learning (Gomes, 2010a; Gomes &Borges, 2009a; Pereira *et al*., 2019; Pires & Gomes, 2018).

The Qualitative Approach is Closer to the Correct Path: The proper evaluation of the student demands that the target construct be estimated based on the own individual. This implies the analysis of the intraindividual variance, which can only be collected if the student performs a task repeatedly times.However, the current quantitative approach to evaluate a student's academic achievement and learning collects data about this student's performance only once as part of a collection of data of many other students. So that, these constructs have their estimates based only on the total collection of students' data. The interindividual variance is the basic data for the production of the estimates. However, interindividual variance is a valuable data when the researcher wants to estimate the construct of the population and therefore, this data is not suitable whenever the researcher wants to make inferences about a student. On the other hand, the intraindividual variance is used in qualitative practices. For example, when an educational psychologist and a teacher argue that a student seems to have an inadequate performance regarding some school content, this inference usually is based on the observation of a variety of behaviors of this student concerning her(his) learning and achievement in this content. In other words, it is very uncommon that an educational psychologist or a teacher would argue that a student shows difficulty in learning some school content by only observing that student once. This qualitative practice is closer to the correct path than the current quantitative approach of the student's evaluation. Since the mainstream practice of quantitative approach is supported by the assumption that the estimate based on population can be

directly transposed to the individual, most of the information is based on the interindividual variance. All the well-known quantitative techniques were created to deal with the interindividual variance, assuming that data are independent, that is, their results are not autocorrelated. Ironically, areas with less influence of the quantitative approach tend to be closer to the right quantitative evaluation of the individual. This occurs specially in some clinical areas that emphasize the processes related to the development of the individual and her(his) progress. A good example is music therapy. This field has many scales that permit the professional to assess the improvement of the patient, as well the progress in therapeutics (André*et al*., 2016). Furthermore, this area has a considerable concern with the process of testing, since many patients have strong cognitive impairment and the process of evaluation is embedded on the own therapeutic process (Rosário *et al*., 2019).

Proper Quantitative Approaches: Following the result of the Birkhoff (1931) Ergodic Theorem the current paradigm on how to evaluate performance of students has to be changed. The evaluation by comparing just one testing result with a population average can no longer be applied. The new paradigm should be done by finding each student average and find the measure of performance by using that average. In our point of view, this new way has yet to be applied in every day psychological testing because this kind of study is very time consuming, very expensive and difficult to carry out. Instead of asking a hundred students to respond a testing quiz once each one, the researcher will have to ask each student to respond the quiz one hundred times. New statistical methodology has to be used like Hidden Markov Models (Gomes &Jelihovschi, 2016), Dynamic Factorial Analysis and Dynamic Item Response Theory (Jelihovschi&Gomes, 2019), which are not yet part of the mainstream methods used by most psychologists. The most difficult part of this new way is the amount of testing each student has to respond, in practice one hundred replications of a test quiz makes it unfeasible to carry out. Jelihovschi and Gomes (2019) tried to find a way out of it by devising a computer program to emulate most of the replications. The program is a function written in R language called*simerg*. This function is described in Jelihovschi and Gomes (2019). In short, the student has to respond only six times the testing procedure. The results are fed to the function which simulates 84 more replications, to complete the desired 90 replications. The results are promising but still more studies and simulations must be done in order to get a scientific acceptable result.

CONCLUSION

In this paper we bring solid arguments indicating that the direct transposition of the population parameters to evaluate the student's parameter is misleading and must be eliminated as a method of evaluation. This current and secular error has innumerable and serious implications, for instance, the lack of confidence on the validity and reliability of the assessment of the student. Validity is also an issue which raises serious doubt about the admission methods in many institutions which use some measurement of the student. If those measurements are not valid, then the admission process is incorrect (Edwards *et al*., 2012). The same is true for the assessment of the student when evaluating his learning and acquisition of knowledge. Imagine the enormous ammount of students in the world which have failed their school grades based on erroneous evaluations. When we apply any exam to a group of students just once, we are able to estimate the parameters of this group and the score of each individual from this type of estimation is a mere stochastic point of a random phenomenon. Imagine that a student performs some famous large-scale assessment, such as PISA. The current approach involves evaluating this student taking the PISA exam only once and then it estimates the parameters of the population which this student is part. In this process, all the students of the population have a score, usually produced by a latent variable approach, such as the item response theory. To believe that this score tells anything about each student of the population is a big mistake. These scores only tell about the random variation of the estimated population.

Each score is a random point. It is not information about each individual since it is not an estimation of any student. The educational system is guided by the stunning error of conception that is present in the secular practice of directly transposing the parameters of population to each individual of the population. The educational system does not know that the scores produced from this rationale tell only about the population or groups since the educational system erroneously understands those scores, which are only random points of the population. This conceptual error should not persist. In sum, considering the arguments presented above, we need, urgently, first recognize that the secular practice of inference about student's achievement is wrong. If we ignore this fact, we will close our eyes to the truth. This paper highlights this embarrassment, claiming that educators should consider this important issue and researchers should think about alternative methodology to make valid and viable the measurement of student's achievement and learning. alternative methodology to make valid and viable the measurement of student's achievement and learning.

REFERENCES

André, A. M., Gomes, C. M. A., & Loureiro, C. M. V. (2016). Escalas Nordoff Robbins: uma revisão bibliográfica [Nordoff Robbins scale: A bibliography review]. *Percepta*, *3*(2), 117-131. https://doi.org/10.34018/2318-891X.3(2)117-131

Birkhoff, G. D. (1931). Proof of the ergodic theorem. Proceedings of the National Academy of Sciences of the United States of America (PNAS),17(12), 656-660. https://doi.org/10.1073/ pnas.17.2.656.

Edwards, D., H. Coates; T. Friedman. (2012). A survey of international practice in university admissions testing. *Higher Education Management and Policy,24* (1). https://doi.org/10.1 787/hemp-24-5k9bdck3bkr8.

Gomes, C. M. A. (2010a). Avaliando a avaliação escolar: notasescolares e inteligênciafluida [Evaluating the school evaluation: the grade schools and fluid intelligence]. *Psicologia em Estudo, 15*(4), 841-849. http://www.redalyc.org/ articulo.oa?i d=287123084020

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

Gomes, C. M. A. (2010c). Perfis de Estudantes e a relação entre abordagens de aprendizagem e rendimento Escolar [Students' Profiles andtheRelationshipbetween Learning Approach andAchievement]. *Psico (PUCRS. Online), 41*(4), 503-509. http://revistaseletronicas.pucrs.br/ojs/index.php/revistapsico/articl e/view/6336

Gomes, C. M. A. (2011a). Abordagem profunda e abordagem superficial à aprendizagem: diferentes perspectivas do rendimento escolar [Deepand surface approach tolearning: different perspectives aboutacademicachievement]. *Psicologia: Reflexão e Crítica*, *24*(3), 438-447. https://doi.org/10.1590/S0102-79722011000300004

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

Gomes, C. M. A. (2012a). A estrutura fatorial do inventário de características da personalidade [The factor structure of the personal characteristics inventory]. *Estudos de Psicologia (Campinas)*, *29*(2), 209-220. https://doi.org/10.1590/S0103-166X2012000200007

Gomes, C. M. A. (2012b). Validade de construto do conjunto de testes de inteligência cristalizada (CTIC) da bateria de fatores cognitivos de alta-ordem (BaFaCAlO) [Constructvalidityofthe set ofcrystallizedintelligencetestsfromhigher-ordercognitivefactors kit]. *Gerais: Revista Interinstitucional de Psicologia*, *5*(2), 294-316.

http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S198 3-82202012000200009&lng=pt&tlng=pt.

Gomes, C. M. A. (2013). A Construção de uma Medida em Abordagens de Aprendizagem [The Constructionof a Measureof Learning Approaches]. *Psico (PUCRS. Online), 44*(2), 193-203. http://revistaseletronicas.pucrs.br/ojs/index.php/revistapsico/articl e/view/11371

Gomes, C. M. A., & Almeida, L. S. (2017). Advocating the broad use of the decision tree method in education. *Practical Assessment, Research & Evaluation, 22*(10), 1-10, 2017. https://doi.org/ 10.7275/y36w-hg55

Gomes, C.M.A., Amantes, A., &Jelihovschi, E.G. (2020). Applying the regression tree method to predict students' science achievement. *Trends in Psychology*, *28*, 99-117. https://doi.org/ 10.9788/s43076-019-00002-5

Gomes, C. M. A., Araujo, J., &Jelihovschi, E. G. (2020). Approaches to learning in the non-academic context: construct validity of learning approaches test in video game (lat-video game). *International Journal of Development Research, 10*(11), 41842-41849. https://doi.org/10.37118/ijdr.20350.11.2020

Gomes, C. M. A., Araujo, J., Nascimento, E., &Jelihovschi, E. (2018). Routine Psychological Testing of the Individual Is Not Valid. *Psychological Reports, 122*(4), 1576-1593. https://doi.org/ 10.1177/0033294118785636

Gomes, C. M. A., & Borges, O. N. (2007). Validação do modelo de inteligência de Carroll em uma amostra brasileira [Validationof Carroll intelligence model in onebrazilian sample]. *Avaliação Psicológica*, *6*(2), 167-179. http://pepsic.bvsalud.org/scielo. php?script=sci_a rttext&pid=S1677-04712007000200007& lng=en&tlng=pt

Gomes, C. M. A., & Borges, O. N. (2008a). Avaliação da validade e fidedignidade do instrumento crenças de estudantes sobre ensino-aprendizagem (CrEA) [Psychometrical proprieties of teaching and learning student'sbeliefsscale]. *Ciências & Cognição (UFRJ), 13*(3), 37-50. http://cienciasecognicao.org/revista/index.ph p/cec/artic le/view/60

Gomes, C. M. A., & Borges, O. (2008b). Qualidades psicométricas de um conjunto de 45 testes cognitivos [Psychometricproprietiesof a set of 45 cognitivetests]. *Fractal: Revista de Psicologia*, *20*(1), 195-207. https://doi.org/10.1590/S1984-02922008000100019

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

Gomes, C. M. A., & Borges, O. N. (2009b). Propriedades psicométricas do conjunto de testes da habilidade visuo espacial [Psychometricproprieties of visual-spatialabilitytests kit]. *PsicoUSF*, *14*(1), 19-34. http://pepsic.bvsalud.org/scielo .php?script= sci_arttext&pid=S1413-8271200900 0100004&ln =pt&tln g=pt

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

Gomes, C. M. A., & Gjikuria, J. (2017). Comparing the ESEM and CFA approaches to analyze the Big Five factors. *Avaliação Psicológica*, *16*(3), 261-267. http://doi.org/10.15689/ap .2017. 1603.12118

Gomes, C. M. A., & Gjikuria, E. (2018). Structural Validity of the School Aspirations Questionnaire (SAQ). *Psicologia: Teoria e Pesquisa*, *34*, e3438. http://doi.org/10.1590/0102.3772e3438

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 [Analysisofcompositetrust scores of enemby facility analysis of items]. *European Journal of Education Studies, 5*(8), 331-344. http://doi.org/10.528 1/zenodo.2527904

Gomes, C. M. A., &Jelihovschi, E. (2016). Proposing a new approach and a rigorous cut-off value for identifying precognition. *Measurement, 93*, 117-125. https://doi.org/10.1016/j. measurement.2016.06.066

Gomes, C. M. A., &Jelihovschi, E. (2019). Presenting the regression tree method and its application in a large-scale educational dataset. *International Journal of Research & Method in Education*, *43*(2), 201-221. https://doi.org/10.1080/ 1743727X. 2019.1654992

Gomes, C. M. A., Lemos, G. C., &Jelihovschi, E. G. (2020). Comparing the predictive power of the CART and CTREE algorithms. *Avaliação Psicológica, 19*(1), 87-96. https://doi.org/10.15689/ap.2020.1901.17737.10

Jelihovschi, E. G., & Gomes, C. M. A. (2019). Proposing an achievement simulation methodology to allow the estimation of individual in clinical testing context. *Revista Brasileira de Biometria, 37*(4), 1-10. https://doi.org/10.28951/rbb.v37i4.423

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2* (4), 201-218, 2004. https://doi.org/10.1207/s15366359mea0204_1

Pereira, B. L. S., Golino, M. T. S., & Gomes, C. M. A. (2019). Investigando os efeitos do Programa de Enriquecimento Instrumental Básico em um estudo de caso único [The effectsoftheFeuerstein's Instrumental Enrichment Basic Program in a single case design]. *European Journal of Education Studies, 6*(7), 35-52. https://doi.org/10.5281/zenodo.3477577

Pires, A. A. M., & Gomes, C. M. A. (2018). Proposing a method to create metacognitive school exams. *European Journal of Education Studies, 5*(8), 119-142. https://doi.org/10.5281/ zenodo.2313538

Rosário, V. M., Gomes, C. M. A., & Loureiro, C. M. V. (2019). Systematic review of attention testing in allegedly "untestable" populations. *International Journal of Psychological Research and Reviews*, *2*(19), 1-21. https://doi.org/10.28933/ijprr-2019-07-1905

****