



EXAMEN

Política, Gestão e Avaliação da Educação



Cebraspe

N. 2

V.2 • jan. - jun. • 2018
ISSN 2526 - 9259

EXAMEN

Política, Gestão e Avaliação da Educação

Revista semestral, volume 2, número 2, janeiro – junho 2018

ISSN 2526-9259

Cebraspe
Brasília, DF – Brasil

Cebraspe

Centro Brasileiro de Pesquisa em Avaliação e Seleção e de Promoção de Eventos (Cebraspe), pessoa jurídica de direito privado, sem fins lucrativos, qualificado por meio do Decreto nº 8.078/2013 como Organização Social (OS), supervisionado pelo Ministério da Educação (MEC), mediante contrato de gestão, com a interveniência da Fundação Universidade de Brasília (FUB) e do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), tendo como finalidade precípua fomentar e promover o ensino, a pesquisa científica, o desenvolvimento tecnológico, o desenvolvimento institucional e a difusão de informações, experiências e projetos de interesse social e utilidade pública nas áreas de avaliação, certificação e seleção.

CONSELHO DE ADMINISTRAÇÃO

ÓRGÃO/ENTIDADE	CONSELHEIRO(A)	
Ministério da Educação	Cleunice Matos Rehem Daniela Helena Oliveira Godoy	Titular Suplente
Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC)	Luiz Fernando Fauth Caroline Menicucci Salgado	Titular Suplente
Instituto Federal de Brasília (IFB)	Wilson Conciani Adilson César de Araújo	Titular Suplente
Rede Nacional de Ensino e Pesquisa (RNP)	Nelson Simões da Silva Wilson Biancardi Coury	Titular Suplente
Associação Brasileira de Estatística (ABE)	Antonio Eduardo Gomes Hélio dos Santos Migon	Titular Suplente
Universidade de Brasília (UnB)	Wellington Lourenço de Almeida Carmenísia Jacobina Aires	Titular Suplente
Universidade de Brasília Conselho Universitário	Germana Henriques Pereira de Sousa Rozana Reigota Naves	Titular Suplente
Universidade de Brasília Conselho de Administração	Fernando César Lima Leite Rivany Borges Beú Bernardes	Titular Suplente
Universidade de Brasília Conselho de Ensino, Pesquisa e Extensão	Adriana Moreira Amado Roberto Goulart Menezes	Titular Suplente
Associação dos Aposentados da FUB	Hildebrando de Miranda Flor Luiz Hernan Rodrigues Castro	Titular Suplente
Representante dos Associados Cebraspe	Noraí Romeu Rocco Marcelo Ladeira	Titular Suplente

DIRETORA-GERAL

Adriana Rígon Weska

DIRETOR EXECUTIVO

José Edil Benedito

DIRETOR DE CONTRATAÇÃO E GESTÃO DE EVENTOS

José Carlos Freitas

DIRETORA DE INSTRUMENTOS DE AVALIAÇÃO, SELEÇÃO E CERTIFICAÇÃO

Claudia Maffini Griboski

DIRETOR DE OPERAÇÕES EM EVENTOS

Jorge Amorim Vaz

CONSELHO EDITORIAL DO CEBRASPE

José Otávio Nogueira Guimarães

Rogério Basali

Denise Aragão

Lucília Garcez

Mauro Luiz Rabelo

PRODUZIDO POR SUPERVISÃO EDITORIAL

SUPERVISORA EDITORIAL Mariana Carvalho

REVISÃO Camila Alves

Danúzia Queiroz

Thaís Lunni

PROJETO GRÁFICO Mariana Bentes

DIAGRAMAÇÃO Rodrigo Carvalho

ENTREVISTA Maíra Andrade

SECRETÁRIA EXECUTIVA DA EXAMEN

Caroline Wollenhaupt Simões Pires

Examen

Examen publica artigos, resenhas e entrevistas que discutem avaliação educacional em larga escala, políticas públicas em educação e certificação educacional e profissional. O propósito da revista é servir como fórum para a apresentação de pesquisas atuais e como veículo de disseminação de informação para comunidade acadêmica, profissionais e sociedade em geral. O periódico é fomentado pelo Cebraspe e é publicado semestralmente.

Artigos, entrevistas e resenhas da publicação não refletem a opinião do Cebraspe ou da revista.

Todos os direitos autorais reservados. É proibida a reprodução integral de artigos.

Examen : pesquisa em avaliação, certificação e seleção / Centro Brasileiro de Pesquisa em Avaliação e Seleção e de Promoção de Eventos – Ano 2, n. 2 (nov. 2018).

Semestral

ISSN 2526-9259

1. Educação – Periódico. 1. Brasil. 2. Educação – Políticas Públicas. 3. Avaliação Educacional.

CDU 371.26(81)

EDITORAS CHEFES

Girlene Ribeiro de Jesus

Universidade de Brasília

Centro Brasileiro de Pesquisa em Avaliação e Seleção e de Promoção de Eventos (Cebraspe)

Brasília, DF – Brasil

Catarina de Almeida Santos

Universidade de Brasília

Brasília, DF – Brasil

EDITORES ASSOCIADOS

Renata Manuely de Lima Rego

Cebraspe

Brasília, DF – Brasil

Letícia Alves Santos

Cebraspe

Brasília, DF – Brasil

Víctor Vasconcelos de Souza

Cebraspe

Brasília, DF – Brasil

CONSELHO CONSULTIVO

Dalton Francisco de Andrade

Universidade Federal de Santa Catarina

Florianópolis, SC – Brasil

Joaquim José Soares Neto,

Universidade de Brasília

Brasília, DF – Brasil

CONSELHO EDITORIAL DA EXAMEN

Cecília Brito Alves

Medical Council of Canada – Canadá

Stephen G. Sireci,

Sireci Psychometric Services

Massachusetts – EUA

José Vieira de Sousa

Universidade de Brasília

Brasília, DF – Brasil

Josemberg Moura de Andrade

Universidade de Brasília

Brasília, DF – Brasil

Alessandra Ramos de Oliveira Harden

Universidade de Brasília

Brasília, DF – Brasil

Felipe Valentini

Universidade Salgado de Oliveira

Niterói, RJ - Brasil

Elizabeth Nascimento

Universidade Federal de Minas Gerais

Belo Horizonte, MG - Brasil

Patrícia Vieira Nunes Gomes

Instituto Nacional de Estudos e Pesquisas

Educacionais Anísio Teixeira (Inep)

Brasília, DF – Brasil

Éverson Meireles

Universidade Federal do Recôncavo da Bahia

Santo Antônio de Jesus, BA – Brasil

Mauro Luiz Rabelo

Universidade de Brasília

Brasília, DF – Brasil

Ricardo Primi

Universidade São Francisco

Itatiba, SP – Brasil

Adriano Borgatto

Universidade Federal de Santa Catarina

Florianópolis, SC – Brasil

Bruno Figueiredo Damásio

Universidade Federal do Rio de Janeiro

Rio de Janeiro, RJ – Brasil

Cláudio Simon Hutz

Universidade Federal do Rio Grande do Sul

Porto Alegre, RS – Brasil

Camila Akemi Karino

Geekie Educacional

São Paulo, SP – Brasil

SU MÁRIO

contents

8 **Apresentação**

Adriana Rigon Weska

10 **Editorial**

Girlene Ribeiro de Jesus e Catarina de Almeida Santos

13 **Artigos**

articles

14 Estudos sobre avaliações em larga escala: um olhar pela análise de redes sociais

Large-scale assessment studies: a view through social network analysis

Estudios sobre evaluaciones a gran escala: una mirada por el análisis de redes sociales

Adriana Bauer

49 Integração leitura/escrita na prova de redação do Vestibular Unicamp: ameaças ao construto?

Reading/writing integration in the Unicamp entrance examination: threats to the construct?

Integración lectura/escritura en la prueba de redacción del examen de ingreso de Unicamp: amenazas al constructo?

Matilde Virginia Ricardi Scaramucci eLuciana Amgarten Quitzau

77 Autoavaliação institucional na política do Sinaes: concepção, desafios e possibilidades de avanços

Institutional self-evaluation in the Sinaes policy: conception, challenges and possibilities of advances

Autoevaluación institucional en la política del Sinaes: concepción, desafíos y posibilidades de avances

José Vieira de Sousa

- 100** Dinâmicas de evasão na educação superior brasileira
Dynamics of drop-out in Brazilian higher education
Dinámicas de evasión en la formación superior brasileña
Luciana Guedes da Silva, Cecília Brito Alves, Joaquim José Soares Neto e Silene de Paulino Lozzi
- 128** A validade do Enade para avaliação de cursos superiores: uma abordagem multinível
The validity of Enade to assess the quality of the undergraduate courses: a multilevel perspective
La validez del Enade para la evaluación de cursos superiores: un enfoque multinivel
Ricardo Primi, Marjorie Cristina Rocha da Silva e Daniel Bartholomeu
- 152** Um estudo da relação entre traço latente e variáveis contextuais no Saeb e Enem
A study relating latent traits to contextual variables for the Saeb and Enem exams
Un estudio de la relación entre trazo latente y variables contextuales en los exámenes Saeb y Enem
Flávio Bambilra Gonçalves e Bárbara da Costa Campos Dias

173 Resenha

review

174 Padrões para educação e psicologia

Standards for educational and
psychological testing

Cristiane Faiad de Moura e Carlos Manoel
Lopes Rodrigues

178 Entrevista

interview

179 O papel da psicometria na testagem em educação e psicologia

The role of psychometry in the test in
education and psychology

El papel de la psicometría en el ensayo
de educación y psicología

Professor Luiz Pasquali

**APRE
SEN
TA
ÇÃO**

contents

O filósofo romano Sêneca é autor de uma frase que resume a importância da comunicação para a nossa sociedade: “Se me oferecessem a sabedoria com a condição de guardar só para mim, sem comunicar a alguém, não a quereria”, escreveu, ainda no século I. Muito tempo se passou desde então, mas a ideia por trás da citação permanece atual. O conhecimento precisa ser socializado, especialmente se ele é fruto de pesquisa acadêmica.

É por isso que apresentamos, com orgulho, a segunda edição da *Examen*, a revista de divulgação científica do Cebraspe. Nela, trazemos artigos, resenhas e entrevistas que versam sobre avaliação educacional, políticas públicas em educação e processos de certificação e seleção. São textos que aprofundam a reflexão sobre os resultados do trabalho desenvolvido pelo Centro, divulgando informações convergentes e divergentes e, portanto, estimulando o pensamento crítico.

O Cebraspe é uma organização que nasceu e consolidou-se fortemente amparada pela pesquisa e pela inovação. Foi graças à inovação e à qualidade de nossos métodos e técnicas que nos tornamos uma das instituições mais reconhecidas do país para realização de avaliações educacionais, concursos públicos e certificações. Assim sendo, temos o dever de continuar a investir na produção e na disseminação do conhecimento. Fazemos isso por meio da divulgação científica e também com editais de fomento à pesquisa, com a realização de seminários acadêmicos e com o intercâmbio permanente com pesquisadores de todo o país. Nos próximos anos, queremos ampliar a colaboração com universidades e instituições de pesquisa, fortalecendo o caráter científico do Centro.

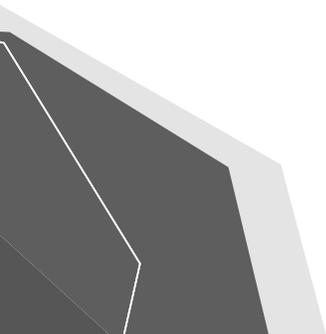
Nas próximas páginas, caro(a) leitor(a), você verá uma amostra da colaboração que o Cebraspe pode dar para a inovação na área educacional. Desejamos que os textos sejam inspiradores e provoquem novas reflexões.

Boa leitura!

Adriana Rigon Weska

EDI TO RIAL

editorial



Temos o prazer de apresentar aos leitores o segundo número da revista *Examen*, periódico científico voltado à disseminação de pesquisas relacionadas à temática de política, gestão e avaliação da educação. É um periódico aberto e todo o conteúdo está disponível livremente. Os trabalhos são publicados e divulgados semestralmente no formato eletrônico. Neste exemplar são oferecidos seis estudos que abordam diversos temas da área de avaliação educacional e políticas públicas em educação, uma resenha da última edição, 2014, do “Standards for Educational and Psychological Testing”, bem como uma entrevista com o professor Luiz Pasquali sobre a temática da validade dos testes.

O primeiro texto desta edição da revista é um “Estudo sobre avaliações em larga escala: um olhar pela análise de redes sociais”, de autoria da professora Adriana Bauer, que retoma resultados de levantamento bibliográfico que analisou a produção acadêmica brasileira sobre avaliações externas e em larga escala no período de 1988 a 2011, discutindo-as a partir da metodologia de análise de redes sociais.

O segundo trabalho, voltado ao estudo de provas de vestibulares, intitulado “Integração leitura/escrita na prova de redação do Vestibular Unicamp: ameaças ao construto?”, de autoria das professoras Matilde Virginia Ricardi Scaramucci e Luciana Amgarten Quitzau, que analisa as ameaças ao construto que podem advir de diferentes propostas de produção escrita, uma vez que essas propostas variam de uma edição para outra em um mesmo exame vestibular.

O professor José Vieira de Sousa assina o terceiro artigo desta edição da revista, intitulado “Autoavaliação institucional na política do Sinaes: concepção, desafios e possibilidades de avanços”. Nesse artigo, reflete-se sobre a autoavaliação institucional como componente do Sistema Nacional de Avaliação da Educação Superior (Sinaes).

O quarto artigo, intitulado “Dinâmicas de evasão no ensino superior brasileiro”, de autoria dos pesquisadores Luciana Guedes da Silva, Cecília Brito Alves, Joaquim José Soares Neto e Silene de Paulino Lozzi, tem como mote principal o seguinte questionamento: é possível antecipar os índices de evasão na educação superior, prevenindo, assim, o sistema das consequentes perdas sociais e econômicas envolvidas na alta evasão?

O quinto artigo, de autoria dos professores Ricardo Primi, Marjorie Cristina Rocha da Silva e Daniel Bartholomeu, intitulado “A validade do Enade para avaliação de cursos superiores: uma abordagem multinível”, tem como proposta identificar qual parcela da variância da prova do Exame Nacional de Desempenho de Estudantes (Enade) está associada ao que se pretende medir e, assim, identificar a sua capacidade de se tornar um indicador da qualidade dos cursos.

O último artigo desta edição, intitulado “Um estudo da relação entre traço latente e variáveis contextuais no Saeb e Enem”, de autoria de Flávio Bamberra Gonçalves e Bárbara da Costa Campos Dias, teve como objetivo aplicar uma metodologia proposta recentemente para estudar a relação entre o traço latente (proficiência) e as variáveis presentes em questionários contextuais em avaliações educacionais de larga escala no Brasil.

Na segunda seção da revista, temos a resenha do livro publicado pela American Educational Research Association (AERA), American Psychological Association (APA) e pela National Council on Measurement in Education (NCME): *Standards for Educational and Psychological Testing*, intitulada: “Padrões para educação e psicologia”.

A revista finaliza com a entrevista “O papel da Psicometria na testagem em educação e psicologia”, que apresenta o ponto de vista do professor Luiz Pasquali - Doutor em Psicologia pela pela Faculté de Psychologie et des Sciences de l'éducat Université Catholique de Louvain, Bélgica, e atualmente professor pesquisador associado à Universidade de Brasília - a respeito da contribuição da Psicometria para avaliação educacional.

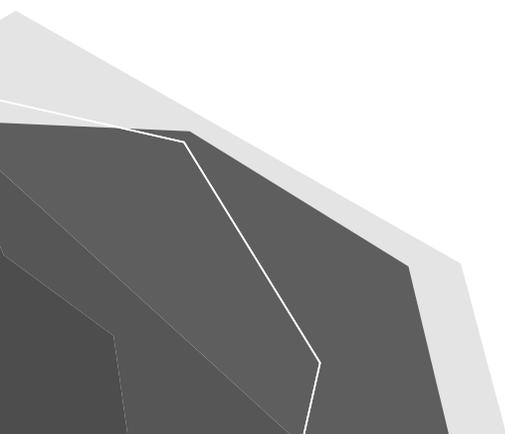
Para concluir, informamos que o primeiro número da *Revista Examen* já está disponível *on-line*. Convidamos você para acessar o portal da revista e desfrutar dos excelentes artigos das duas edições.

Desejamos a todos uma proveitosa leitura!

Girlene Ribeiro de Jesus
Catarina de Almeida Santos

AR TI GOS

articles



ESTUDOS SOBRE AVALIAÇÕES EM LARGA ESCALA: UM OLHAR PELA ANÁLISE DE REDES SOCIAIS¹

LARGE-SCALE ASSESSMENT STUDIES: A VIEW THROUGH SOCIAL NETWORK ANALYSIS

ESTUDIOS SOBRE EVALUACIONES A GRAN ESCALA: UNA MIRADA POR EL ANÁLISIS DE REDES SOCIALES

Adriana Bauer

RESUMO

Este artigo retoma resultados de levantamento bibliográfico, que analisou a produção acadêmica brasileira sobre avaliações externas e em larga escala no período de 1988 a 2011, discutindo-os a partir da metodologia de análise de redes sociais. Tem como objetivo contribuir para: (1) o entendimento das relações que se estabelecem entre atores que têm se dedicado sistematicamente à discussão da temática; e (2) refletir sobre a qualidade dessa produção. Os resultados apontam para uma dispersão na produção e avaliação do conhecimento, sendo que diversos trabalhos são orientados e avaliados, em sua maioria, no âmbito de comunidades isoladas e não parecem dialogar diretamente com a produção mais consolidada sobre a temática. Como possível consequência desse isolamento, discute-se a tendência de enfraquecimento teórico-metodológico e a sobreposição de trabalhos sobre os mesmos aspectos e temas, que dificultam o aprofundamento da discussão e, possivelmente, o avanço no conhecimento já existente.

Palavras-chave: avaliação em larga escala; produção acadêmica; análise de redes sociais; levantamento bibliográfico.

¹ A autora agradece a Eduardo Rodrigues Capocchi pelo processamento das redes sociais discutidas no trabalho e pelas análises estatísticas a elas relativas.

ABSTRACT

This paper reviews results of a bibliographical survey, which aimed to analyze the Brazilian academic production on external and large-scale assessments from 1988 to 2011, discussing them from the social networks analysis methodology. It aims to contribute to: (1) the understanding of the relationships established by actors who have been systematically engaged in the discussion of this theme; and (2) to think about the quality of this scientific production. The results point to a dispersion in the production of knowledge; it also points that several works are produced and evaluated by isolated communities and do not seem to dialogue directly with the more consolidated production on the subject. As a possible consequence of this isolation, we discuss the tendency of theoretical-methodological weakening and the overlapping of works on the same aspects and themes, which make it difficult to deepen the discussion and, possibly, the advance in the already existing knowledge.

Keywords: large-scale assessment; academic production; social network analysis; bibliographic survey.

RESUMEN

Este artículo retoma resultados de levantamiento bibliográfico que objetivó analizar la producción académica brasileña sobre evaluaciones externas y en gran escala en el período de 1988 a 2011, discutiéndolos a partir de la metodología de análisis de redes sociales. Se ha dirigido a contribuir a: (1) la comprensión de las relaciones de los actores que se han dedicado sistemáticamente a la discusión temática; y (2) reflejar acerca de la cualidad de esta producción. Los resultados apuntan a una dispersión en la producción del conocimiento, siendo que diversos trabajos son producidos y evaluados por comunidades aisladas y no parecen dialogar directamente con la producción más consolidada sobre la temática. Como posible consecuencia de este aislamiento, se discute la tendencia de debilitamiento teórico-metodológico y la superposición de trabajos sobre los mismos aspectos y temas, que dificultan la profundización de la discusión y, posiblemente, el avance en el conocimiento ya existente.

Palabras clave: evaluación a gran escala; producción académica; análisis de redes sociales; levantamiento bibliográfico.

Introdução

Tem-se observado, no Brasil, aumento substantivo dos cursos de pós-graduação. Dados divulgados pela Geocapes para 2012 atestaram a existência de 1.664 programas de mestrado e doutorado, 1.230 programas somente de mestrado, 395 programas de mestrado profissional e 53 programas de doutorado. Tal aumento associa-se, ainda que não necessariamente de forma linear, ao crescimento da produção científica nas diferentes áreas e temáticas, quer por meio da divulgação das teses e dissertações, quer por meio de sua difusão em forma de artigos.

Pesquisa de Santos e Farias (2010) ilustra essa tendência ao mostrar que a produção científica brasileira subiu 56% entre 2007 e 2008, passando de 19.436 artigos para 30.451. No entanto, ainda que o Brasil tenha subido do 17º para o 13º lugar no *ranking* mundial de quantidade de artigos publicados, “a qualidade dos trabalhos científicos, medida, por exemplo, pelo número de vezes que cada estudo foi citado por outros cientistas (o chamado ‘impacto’) despencou. O Brasil passou de 31º lugar para 40º” (RIGHETTI, 2013). Mesmo que as discussões sobre medição da qualidade da produção sejam controversas e acaloradas, pode-se dizer que, **ainda antes** da ênfase na produção de métricas que aquilatassem a qualidade dos trabalhos acadêmicos, os clamores acerca da necessidade de avaliação da qualidade dessa produção já existiam.

Desde os anos de 1980, diversos autores já advertiam para o imperativo de se buscar rigor e qualidade na pesquisa (ANDRÉ, 2001), de se formular objetivos de pesquisa relevantes na pós-graduação (LAROSSA; ROSSO; SOUZA, 2005), de questionar os instrumentos teórico-metodológicos utilizados e os parâmetros existentes para avaliação da qualidade da pesquisa científica (GATTI, 2001), de se atentar para a fragilidade da revisão bibliográfica nas teses e dissertações (ALVES, 1992) e para aspectos de relevância e aplicabilidade da pesquisa (ALVES-MAZZOTTI, 2001). Alguns dos problemas apontados por esses autores, ainda, parecem bastante atuais.² Claro que não se pode desconsiderar nas

² Pesquisa bibliográfica que gerou a base de dados utilizada neste artigo permitiu perceber, nos trabalhos que compõem seu *corpus*, certa pobreza teórico-metodológica, mais acentuada em relação aos mestrados, marcada por estudos descritivos e exploratórios, cuja contribuição ao conhecimento existente no campo não fica clara. Percebe-se, ainda, pouco aprofundamento em torno de temas candentes, pouco cuidado na seleção de referenciais teórico-metodológicos e na divulgação restrita dos resultados que, além de tudo, não impac-

análises os processos de pesquisa que produzem seus resultados. Na área educacional, a pesquisa é cada vez mais realizada de forma isolada e sem condições ideais de produção, como a dedicação exclusiva aos estudos. No entanto, ainda que tais variáveis³ possam explicar o fenômeno, o problema da qualidade dos trabalhos continua posto.

Neste texto, reconhece-se a insuficiência dos mecanismos de avaliação externa existentes (Avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Capes) para aquilatar a qualidade e relevância dessa produção em termos globais; no entanto, a discussão proposta direciona-se ao entendimento de processos internos e aos programas que visam a garantia dessa qualidade, mais especificamente, a constituição da banca para o exame dos trabalhos no momento de sua defesa. Assim, este estudo detém-se na análise das conexões e interações existentes entre aqueles que se responsabilizam, direta e indiretamente, pela qualidade dos trabalhos defendidos nos programas de pós-graduação.

O *corpus* de teses e dissertações escolhido para basear as análises trata da temática da avaliação externa e em larga escala, dimensão específica da avaliação educacional. Os trabalhos analisados foram pesquisados no banco de teses e dissertações da Capes e cobrem o período que vai de 1988 a 2011. As análises norteiam-se pelo conceito do campo científico de Bourdieu (2003)⁴ e da comunidade epistêmica de Peter Haas (1992), que respaldam a utilização da metodologia de análise de redes sociais (*social network analysis*).

Entende-se que a compreensão do campo de estudos sobre avaliação educacional – e da dimensão de avaliação externa e em larga escala mais especificamente –, leva à necessidade de se considerar os sujeitos que se dedicam a investigar a temática e as relações que se estabelecem entre eles, partindo-se do princípio de que tais vínculos assumem significados distintos e influenciam nos resultados obtidos.

tam sobre as práticas analisadas. Tais aspectos reforçam a apreciação de Alves-Mazzotti, do início dos anos 2000, sobre a qualidade da pesquisa em educação (ALVES-MAZZOTTI, 2001, p. 40).

3 Investigações epistemológicas poderiam acrescentar outros aspectos a esses já conhecidos.

4 Há bastante debate acerca da natureza da avaliação educacional: se ela constituiu um campo de conhecimento específico (DIAS SOBRINHO, 2002), se ela é uma disciplina ou uma transdisciplina (SCRIVEN, 2008). Nos limites deste trabalho, não é possível estender o argumento, mas ela será tratada como campo de conhecimentos.

Parte-se da hipótese de que o fato da temática em tela⁵ ter se tornado uma temática “da moda”, aliado à facilidade de acesso aos dados das avaliações existentes e a natureza transdisciplinar da avaliação educacional, tem contribuído para o aumento da produção acadêmica, sem que ele seja acompanhado, necessariamente, da melhoria da qualidade dos estudos. Antes, parece que é um conhecimento produzido por comunidades com interesses pontuais na temática, não permitindo aprofundamento do conhecimento já acumulado.

Método

Como dito anteriormente, o estudo baseia-se em teses e dissertações pesquisadas no Banco de Teses e Dissertações da Capes, organizadas com auxílio da ferramenta Access, em um banco de dados que gerou as informações aqui discutidas.⁶ A construção das redes sociais exploradas no artigo foi realizada por meio do pacote *igraph* (CSARDI; NEPUSZ, 2006) no ambiente computacional R.⁷

Eduardo Marques (1999, p. 46) define rede social como “campo, presente em determinado momento, estruturado por vínculos entre indivíduos, grupos e organizações construídos ao longo do tempo”. Tais vínculos produzem e reproduzem relações que são definidoras de uma rede social. Para Fialho (2014, p. 11), a distinção conceitual entre “redes sociais” e “redes espontâneas e naturais” está na “intencionalidade dos relacionamentos e nos objetivos comuns estabelecidos entre os elementos que nelas (redes) interagem”.

Estudar padrões de colaboração científica ou de constituição de campo de conhecimento por meio da Análise de Redes Sociais (ARS) não é exatamente uma nova abordagem. Diversos trabalhos têm utilizado essa perspectiva

5 Avaliação externa e em larga escala.

6 Para maiores informações sobre o estudo e acesso às análises descritivas que o compõem, consultar Bauer (2012 e 2014); e Bauer e Reis (2013).

7 Referências práticas para o estudo de redes sociais a partir do ambiente R foram pesquisadas em Conway (2009). Outros estudos foram utilizados como referencial teórico para o entendimento da abordagem, tais como: Marques (1999), Borgatti et al. (2013), Fialho (2014), Molina et al. (2005), Penna Neto; Frey e Czajkowski Jr. (2007), Marteleto (2001) e Soczka (2001).

analítica (MARTINS, 2012; SILVA, 2013; BORDIN; GONÇALVES; TODESCO, 2014; SACERDOTE, 2018), ainda que em outros campos científicos e para investigar outros temas.

A análise dessas redes contribui para a localização dos atores relevantes na constituição de um campo. A abordagem constitui, portanto, uma ferramenta que pode ser utilizada em diversas situações em que é necessário analisar e interpretar relações que se estabelecem entre atores. Para o estudo aqui definido, importa identificar os atores que mais se destacam e que, potencialmente, podem influenciar ou não os demais grupos.

Não se desconsidera que as relações e interações que se estabelecem entre os pesquisadores podem ser marcadas por vários fatores, para além da sua *expertise* na temática do trabalho a ser avaliado, tais como: laços de amizade, disponibilidade de agenda e de recursos e conveniência geográfica ou institucional, entre outros. No entanto, concorda-se com Martins quando afirma que esse tipo de mapeamento:

[...] pode fornecer dados relevantes sobre várias camadas em que essas relações ocorrem: estratégias de conectividade utilizadas, políticas de relação, ações recorrentes denotando padrões em potencial, linguagem e expressões características, objetos de pesquisa preferenciais, modos de relação entre instituições, departamento, grupos de pesquisa, pesquisadores, professores e alunos (MARTINS, 2012, p. 3).

Para melhor compreensão dos resultados, cabe destacar alguns conceitos subjacentes às análises realizadas. Recorre-se aos trabalhos de Borgatti, Everett e Johnson (2013), Souza e Quandt (2008) e Fialho (2014) para sistematizar esses conceitos.

Atores ou nós: são as entidades (pessoas ou instituições, cidades ou países, por exemplo) que formam um sistema. Os nós têm seus atributos, que podem ser qualitativos ou quantitativos.

Vínculos, laços ou links: estabelecem relações entre pares de autores e caracterizam as relações que se estabelecem entre eles. Por exemplo, a

relação entre dois atores, um homem e uma mulher, pode ser de casamento, de trabalho, de estudo, etc. Em uma análise de rede, observa-se que os *links* que se estabelecem podem interligar diferentes atores ($A \rightarrow B$; $B \rightarrow C$, sendo que B é um nó que se interliga a A e a C).

Caminho: sequência de nós adjacentes.

Grafo de rede ou sociograma: composto por nós e *links*.

Centralidade da rede: medida de quanto um ator está conectado a outro; neste estudo, indica-se o quanto um pesquisador está conectado a outros membros da comunidade científica, com indicação de sua proeminência e influência sobre o grupo.

Além da análise das redes sociais produzidas para este estudo, são retomadas características do referencial teórico dos trabalhos já descritas em outros estudos (BAUER, 2012 e 2014; REIS, 2013).

Resultados e discussão

A primeira análise expressa os relacionamentos entre os orientadores e os membros de banca⁸ estabelecidos nos 294 trabalhos levantados sobre avaliação externa e em larga escala produzidos entre 1988 e 2011. Ao todo, foram mapeados 676 nomes de acadêmicos envolvidos na avaliação dessa produção.

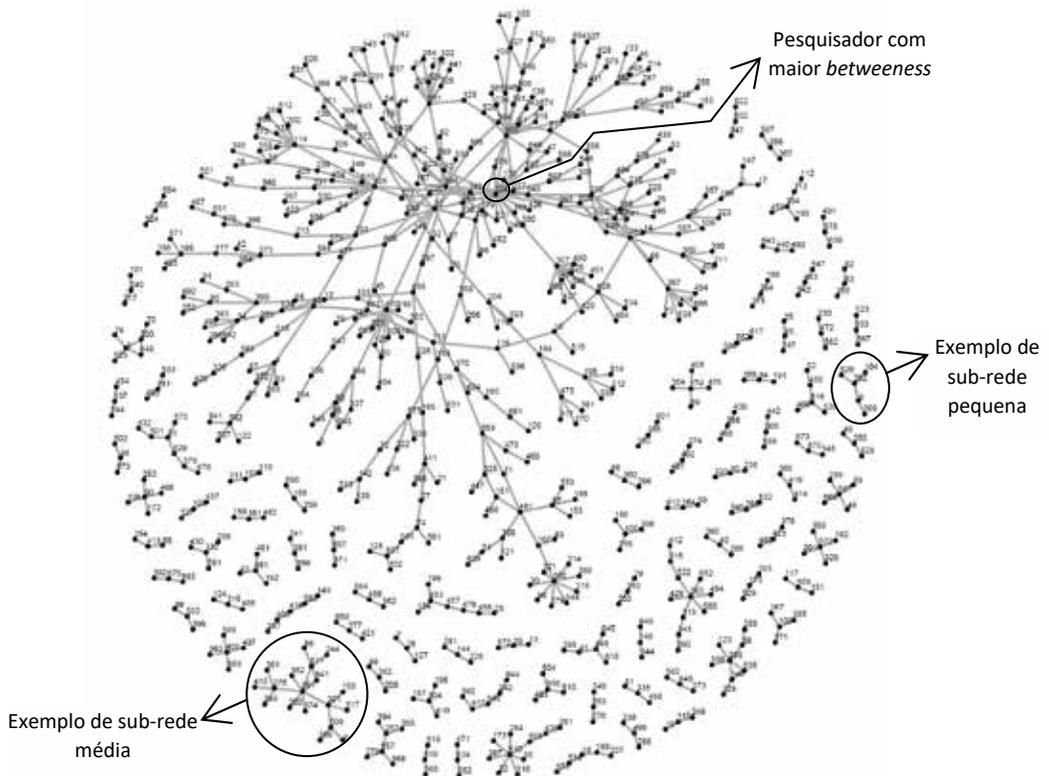
A figura 1⁹ representa o grafo completo de rede. Cada nó representa um indivíduo, que pode ter atuado como orientador ou como membro da banca examinadora de um ou mais trabalhos. As ligações entre os nós representam uma conexão social na qual os indivíduos representados por cada nó colaboraram.¹⁰

8 As bancas aqui referidas são as de conclusão do trabalho.

9 No pacote *igraph* no R, utilizou-se o algoritmo Fruchterman-Reingold com 50 mil iterações para estabelecer o *layout* dos sociogramas.

10 Não foram usados ponderadores para estabelecer colaborações repetidas em distintos trabalhos na construção do grafo de rede (figura 1). Assim, mesmo que indivíduos tenham colaborado em mais de um trabalho, apenas uma aresta é representada. Além disso, não se

Figura 1 Sociograma elaborado com informações sobre os membros das bancas dos trabalhos sobre avaliação em larga escala defendidos entre 1988 e 2011



Observa-se que o sociograma é composto por uma rede maior, acompanhada por redes médias e redes menores, isoladas, com três a cinco colaboradores cada. A sub-rede maior reúne 348 dos 676 acadêmicos e interliga mais da metade dos indivíduos (51,4%), é nela que se pode observar o maior número de interações entre indivíduos e distinguir indivíduos mais centrais, a partir dos quais se estabelecem várias conexões. Há outros 73 acadêmicos que fazem parte de oito sub-redes com 6 a 18

estabeleceu o fluxo das relações (se unilateral ou bilateral), pois, dada a quantidade de nós, seria difícil analisar esses fluxos, mesmo com a imagem ampliada.

nomes cada (10,8%). Finalmente, há 75 sub-redes isoladas que envolvem 255 acadêmicos (37,7%). Estas últimas representam teses (redes de quatro e cinco nós) e dissertações (redes de três nós) únicas, defendidas com bancas formadas por orientadores e avaliadores “isolados” das demais **sub-redes**¹¹, sugerindo que a participação do orientador e dos membros da banca, em trabalho do campo, foi pontual.¹²

A tabela 1 mostra a relação dos atores com os maiores graus de conexão, em ordem decrescente, complementando a análise da rede de grafos mostrada anteriormente. A análise do grau de conexões¹³ de cada indivíduo permite afirmar a centralidade do pesquisador Francisco Creso Franco Jr. (representado pelo número 224 no grafo) na configuração do campo. Seu grande número de conexões (29) e sua posição na representação revelam sua interação com diversos outros indivíduos, muitas vezes, sendo a conexão entre representantes de diversas áreas de conhecimento.¹⁴

11 Dalton Lopes Martins (2012, p. 172) obteve resultado similar na análise de teses e dissertações defendidas em 2011, pesquisadas na Base de Teses e Dissertações do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), com outro foco temático. O autor também encontrou, em um *corpus* distinto de trabalhos, a existência de uma sub-rede maior acompanhada de sub-redes menores, algumas das quais isoladas.

12 Os nomes dos acadêmicos utilizados na composição da figura são listados no quadro 2, ao final do trabalho.

13 Tais métricas visam quantificar propriedades das redes e de seus membros quanto à importância de seus papéis relativos. A métrica mais simples aplicável a um nó da rede é seu grau de conexões, ou o número de outros nós com os quais se liga.

14 Outros pesquisadores que se sobressaem pelo seu número de conexões, que representam a quantidade de trabalhos orientados e/ou participação em bancas de conclusão de mestrado e doutorado são: Sandra Zákia Sousa, Tufi Machado Soares, Jacob Laros, Alicia Bonamino, Flávia Werle, Isabel Cappelletti (*in memoriam*), Naércio de Menezes, Elaine Pasello, José Francisco Soares, Robert Verhine, Romualdo Portela de Oliveira e Luis Pasquali, dentre outros.

Tabela 1 Valores de *betweenness* dos orientadores e membros de banca de defesa dos trabalhos sobre avaliação externa defendidos entre 1988 e 2011 representados na figura 1

NOME	NÚMERO DE CONEXÕES	NOME	NÚMERO DE CONEXÕES
Franco_Jr_FCJ	29	Oliveira_RLP	12
Sousa_SZL	23	Pasquali_L	11
Soares_TM	21	Santos_MCd	10
Laros_JA	20	Fernandes_R	9
Bonamino_AMC	17	Brooke_NPd	8
Werle_FOC	15	Oliveira_DA	8
Cappelletti_IF	14	Rigotti_JIR	8
Menezes_Fo_NA	14	Scorzafave_LGD	8
Pazello_ET	14	272 nomes	2 a 7
Soares_JF	14	385 nomes	1
Verhine_RE	14	Total de 676 nomes	

Fonte: base de dados da pesquisa. Elaboração: Eduardo Capocchi

Tal afirmação é respaldada pela métrica *betweenness*, cujo valor denota a importância de um indivíduo (nó) em conectar outros indivíduos ou sub-redes que, sem ele, não seriam ligados. Os 20 pesquisadores com *betweenness* mais altos, e suas respectivas métricas, calculadas pelo *igraph*, são mostrados na tabela 2.

Tabela 2 Valores de *betweenness* dos orientadores e membros de banca de defesa dos trabalhos sobre avaliação externa defendidos entre 1988 e 2011 representados na figura 1

NOME	BETWEENNESS NORMALIZADA	NOME	BETWEENNESS NORMALIZADA
Franco_Jr_FCJ	0.09	Fernandes_R	0.03
Bonamino_AMC	0.09	Pazello_ET	0.03
Soares_JF	0.07	Brooke_NPd	0.03
Vianna_HM	0.07	Almouloud_SA	0.03
Sousa_SZL	0.06	Werle_FOC	0.03
Oliveira_RLP	0.05	Pires_CMC	0.03
Massi_CDB	0.04	Sousa_JVd	0.02
Soares_TM	0.04	Verhine_RE	0.02
Santos_VdM	0.04	238 nomes	0 < bet < 0.02
Freitas_LCd	0.03	418 nomes	Zero
Sordi_MRL	0.03		
Gremaud_AP	0.03	Total de 676 nomes	1

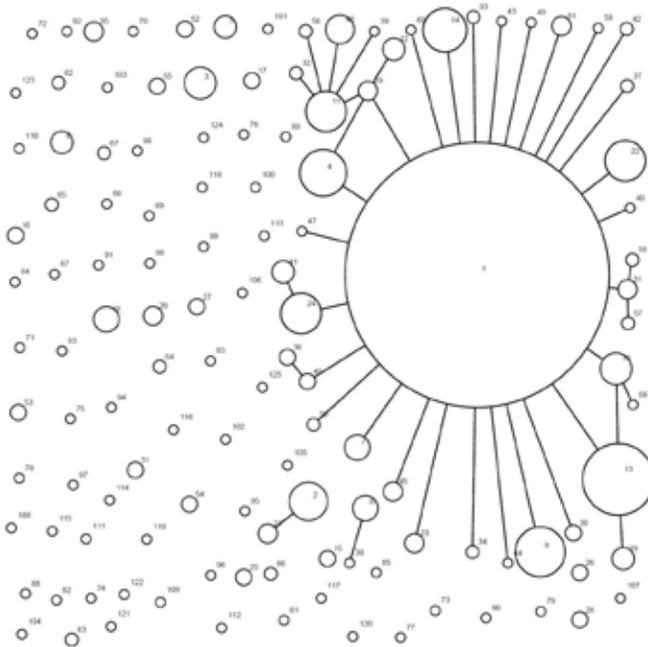
Fonte: base de dados da pesquisa. Elaboração: Eduardo Capocchi.

A ausência permanente ou temporária de um indivíduo com alto grau de *betweenness* tende a enfraquecer a rede como um todo. Por exemplo, a saída do campo de estudos de indivíduos como Heraldo Vianna (*in memoriam*) e Creso Franco, tende a impactar as relações entre os indivíduos e as sub-redes, caso não haja outro pesquisador que assuma o papel articulador dos pesquisadores citados.

Outra forma de compreender as relações que se estabelecem entre os diversos atores, no período estudado, pode ser feita por meio da aplicação do conceito de *comunidade*. A figura 2 representa graficamente¹⁵ as comunidades calculadas a partir da base de dados, sendo que o diâmetro dos círculos representa seu tamanho. Como referência, a maior comunidade, de número 1, tem 85 membros participantes. A segunda maior, de número 13, tem 23 participantes. As menores têm apenas três membros.

¹⁵ As comunidades foram calculadas no *igraph* por meio do algoritmo de *random walks*, que permite observar a existência de 125 agrupamentos aproximados. Isso não significa que os atores que fazem parte de uma mesma comunidade tenham algum tipo de produção colaborativa ou pertençam às mesmas instituições, por exemplo. O que se observa, na análise da composição de cada comunidade, é que a frequência de colaborações entre os membros de cada comunidade é maior.

Figura 2 Representação das comunidades existentes considerando toda a base de dados investigada



Fonte: base de dados da pesquisa. Elaboração: Eduardo Capocchi

O algoritmo utilizado estabeleceu a existência de 125 comunidades, sendo a maioria delas “isoladas”, com uma contribuição muito pontual à orientação de trabalhos e à avaliação da produção acadêmica sobre a temática. A existência de tantas comunidades “isoladas”, que não necessariamente têm domínio sobre o *corpus* acadêmico já existente, poderia explicar a heterogeneidade no uso de terminologias, conceitos e abordagens metodológicas no campo pesquisado, revelada no estudo descritivo (BAUER, 2012 e 2014; BAUER; REIS, 2014).

A tabela 3 relaciona a quantidade de trabalhos orientados e avaliados pelos tipos de comunidades existentes. Observa-se que são as comunidades isoladas que têm se envolvido mais na orientação e avaliação do maior número de trabalhos.

Tabela 3 Quantidade de trabalhos avaliados e orientados por tamanho de comunidade

TAMANHO DA COMUNIDADE ¹⁵	Nº DE TRABALHOS AVALIADOS/ ORIENTADOS
Isolada (3 a 5 membros)	103
Pequena (6 a 9 membros)	39
Média (10 a 33 membros)	71
Grande (85 membros e acima)	81
Total	294

Fonte: base de dados da pesquisa. Elaboração: Eduardo Capocchi

Para compreender os referenciais teóricos mais utilizados pelos autores dos trabalhos, foi feita uma seleção, elegendo-se para análise os 42¹⁷ doutorados que fazem parte da base de dados. Partiu-se das indicações existentes nas referências bibliográficas, desde que citadas no trabalho, selecionando as que continham, no título, as palavras “avaliação”, “avaliação externa”, “avaliação em larga escala”, “avaliador/avaliar”, “mensuração”, “testes padronizados”, “exame”, “política de avaliação”, “reforma”, “desempenho”, “rendimento”, “proficiência”, “indicadores”, “prova”, “monitoramento”, “regulação”, “*accountability*”, “responsabilização”, “efeito-escola” e “qualidade”, e seus equivalentes em inglês, espanhol e francês. Essa seleção baseou-se em léxico comum aos trabalhos analisados, reportando tendência de palavras utilizadas como descritores de busca. Na seleção das referências, consideraram-se, ainda, as denominações dos sistemas estaduais e federais de educação. A seleção resultou em 382 obras de referência, sendo que 100 delas obtiveram mais de uma citação.

Em seguida, selecionaram-se apenas as obras dos autores brasileiros, buscando identificar os mais citados, quer em trabalhos produzidos individualmente, quer em trabalhos colaborativos. Os resultados são expostos nas tabelas 4 e 5.

16 Um dado trabalho pode ser oriundo de mais do que uma comunidade se os acadêmicos associados a ele estiverem em comunidades distintas, isto se deve à natureza randômica do algoritmo de *clustering* das comunidades usado. Quando este é o caso, o tamanho da comunidade usado na quantificação dos trabalhos foi a soma dos tamanhos das comunidades envolvidas.

17 Na base de dados, constam 44 trabalhos de doutoramento. No entanto, só foi possível acessar, eletrônica ou fisicamente, as 42 teses acima referidas.

Tabela 4 Referenciais teóricos produzidos por pesquisadores brasileiros individualmente mais citados nos trabalhos da base de dados

AUTOR	Nº DE CITAÇÕES	AUTOR	Nº DE CITAÇÕES
VIANNA, H. M.	24	ARRETCHE, M.T. da S.	4
SOUSA, S. M. Z. L.	20	BARRETTO, E. S. de S.	4
DIAS SOB, J.	13	HOFFMANN, J. M. L.	4
FRANCO, C.	9	SOUZA, C.P. de	4
SOARES, J. F.	9	CAPPELLETTI, I. F.	3
BONAMINO, A. M. C.	8	DALBEN, A. I. L. de F.	3
GATTI, B. A.	8	FREITAS, D. N. T. de	3
FREITAS, L. C.	7	MACHADO, C.	3
FONTANIVE, N. S.	6	MINHOTO, M. A. P.	3
KLEIN, R.	6	SAUL, A. M.	3
LUCKESI, C. C.	6	SCHWARTZMAN, S.	3
PESTANA, M. I. G. S.	6	SORDI, M. R. L. de	3
SOARES, T. M.	6	-	-
BROOKE, N.	5	-	-
FLETCHER, P. R.	5	14 outros autores	2
MALUF, M. M. B.	5	72 outros autores	1
SOUZA, P. R.	5	-	-

Fonte: base de dados da pesquisa. Elaboração da autora.

Tabela 5 Referenciais teóricos produzidos em colaboração entre dois ou mais pesquisadores mais citados nos trabalhos da base de dados

PRODUÇÃO COLABORATIVA	Nº CITAÇÕES
KLEIN, R.; FONTANIVE, N. S.	5
BONAMINO, A. M. C. de; FRANCO, C.	5
ALBERNAZ, A.; FERREIRA, F. H. G.; FRANCO, C.	4
BARBOSA, M. E. F.; FERNANDES, C.	4
FONTANIVE, N. S.; KLEIN, R.	4
GOMES NT, J. B.; ROSENBERG, L.	4
SOUSA, S. M. Z. L.; OLIVEIRA, R. P. de	4
BONAMINO, A. M. C. de; BESSA, N.; FRANCO, C.	3
FRANCO, C.; FERNANDES, C.; SOARES, J. F.; BELTRÃO, K.; BARBOSA, M. E.; ALVES, M. T. G.	3
SOARES, J. F.; CANDIAN, J. F.	3
12 produções colaborativas	2
42 produções colaborativas	1

Fonte: base de dados da pesquisa. Elaboração: a autora.

Ainda que se perceba claramente, pela análise das tabelas, que alguns autores têm sido mais reconhecidos como referências no campo, considera-se relativamente baixo o número de citações obtidas mesmo por aqueles com melhores posicionamentos nas tabelas. Ao que parece, nem todos os pós-graduandos dialogam com a produção realizada no país já existente na área. Cabe comentar, ainda, que dificilmente há interlocução com os próprios mestrados e doutorados que vêm sendo produzidos ao longo do período de estudos.¹⁸ Não se quer defender, aqui, que todos os pós-graduandos devessem se embasar nos mesmos autores. O que se argumenta é que as análises não evidenciam uma apropriação forte dos estudos efetuados sobre a temática nem por aqueles que a investigam.

¹⁸ Ainda que nos limites deste artigo não seja possível explorar esse dado, cabe esclarecer que, durante a leitura dos 42 trabalhos de doutorado e sua sistematização em ficha própria, procurou-se observar se eram citados trabalhos de mestrado e doutorado defendidos no período. O que se percebeu foi pouca interlocução dos autores dos doutorados com outros trabalhos recém-produzidos em nível de pós-graduação.

Relações que se estabelecem em subgrupos temáticos

A partir da leitura e análise dos títulos e dos resumos dos 294 trabalhos, foi possível categorizá-los em quatro grupos distintos, com base nos focos de interesse identificados nos estudos. Cabe salientar que o exercício de definição destes subgrupos não pode ser considerado exaustivo e ressalta-se que várias das dissertações e teses poderiam ser classificadas em mais de um dos subgrupos.

No grupo 1, foram considerados os estudos que têm como objeto a avaliação externa e em larga escala propriamente dita, propondo reflexões acerca das políticas de avaliação e seu papel na política educacional ou na reforma educacional, seus condicionantes, motivações, pressupostos que a baseiam, delineamentos, desenhos metodológicos, planejamento e execução, etc. O grupo 2 refere-se às pesquisas sobre as implicações dessas avaliações no sistema educacional e na escola. É composto por teses e dissertações que discutem os usos (e desusos) dos resultados das avaliações na gestão, no planejamento escolar, implicações das avaliações no currículo, na dinâmica escolar, na formação e trabalho dos professores ou, ainda, que usam os dados das avaliações para discutir os fatores explicativos dos resultados obtidos pelos alunos. Análises de efeitos e impactos dessas avaliações fazem parte desse agrupamento. O grupo 3 foi engloba os estudos que fazem uma discussão centrada no desenho ou na metodologia da avaliação de sistemas educacionais, nos instrumentos utilizados e no processo de elaboração e análise de conteúdo dos itens, das matrizes de referência, das escalas de proficiência. Estudos cujo foco é a discussão dos procedimentos estatísticos utilizados para processar os dados e/ou a transformação desses dados em indicadores para análise e monitoramento dos resultados educacionais também foram incorporados nesse agrupamento.

Finalmente, o grupo 4 é aquele dos estudos que se apropriam das informações geradas pelas avaliações para fazer discussões e análises a partir de outros objetos de estudo, ou seja, o interesse não está na discussão da avaliação de sistemas educacionais, mas em objetos de estudo específicos da Matemática, da Língua Portuguesa, da Demografia, do Planejamento Urbano e Regional, etc. Também foram identificados estudos cuja preocupação central é discutir as diferenças nos resultados devido à diversidade de gênero, raça e etnia, utilizando as bases de dados das avaliações que, por não se aterem às

discussões de avaliação, foram computados nesse grupo. A tabela 6 relaciona os subgrupos à quantidade de trabalhos produzidos. Informações descritivas adicionais são encontradas na tabela 8, nos anexos do artigo.

Tabela 6 Distribuição dos estudos pelos subgrupos

GRUPOS	FREQUÊNCIA	%
Grupo 1 – Foco nas políticas de avaliação	57	19,4
Grupo 2 – Foco nas implicações das avaliações	102	34,7
Grupo 3 – Foco nos desenhos metodológicos e tratamento dos dados	41	13,9
Grupo 4 – Uso das avaliações em estudos que focalizam outros objetos de análise	94	32,0
Total	294	100,0

Fonte: base de dados da pesquisa. Elaboração da autora.

Optou-se por fazer o grafo para esses subgrupos, partindo da hipótese de que, nos subgrupos temáticos, talvez houvesse um menor número de grupos isolados e mais conexões entre os pesquisadores. Em outras palavras, a aposta foi de que a afinidade nas linhas temáticas pudesse refletir em comunidades epistêmicas mais densas.

As figuras 3 a 6 representam os sociogramas para cada um dos grupos temáticos estudados. Em uma primeira análise, observa-se que se repete a tendência, já presente na figura 1, de se obter uma rede maior com redes menores isoladas. O que varia, no entanto, quando se interpreta as figuras comparativamente, é o número de redes isoladas em cada subgrupo e o tamanho e desenho das redes maiores; além disso, a visualização gráfica permite perceber que, nas temáticas dos grupos 2 e 4, há maior número de redes isoladas, sugerindo maior dispersão em relação aos grupos 1 e 3. Nestes grupos, o percentual de participantes nas redes isoladas é de 54,4% e 58,9%, respectivamente. O grupo com menor percentual de participantes em redes isoladas é o grupo 3 (41,75%) que, também, é o grupo com maior concentração de acadêmicos na rede maior (51,5%). A especificidade da temática que une os participantes do grupo 3 e o seu nível de especialização, que requer conhecimento de metodologias de avaliação e técnicas avançadas de análise de dados, pode ser fator explicativo de sua configuração.

Figura 3 Grafo representativo dos acadêmicos envolvidos nas bancas de defesa de trabalhos do grupo 1, com foco na discussão de políticas de avaliação

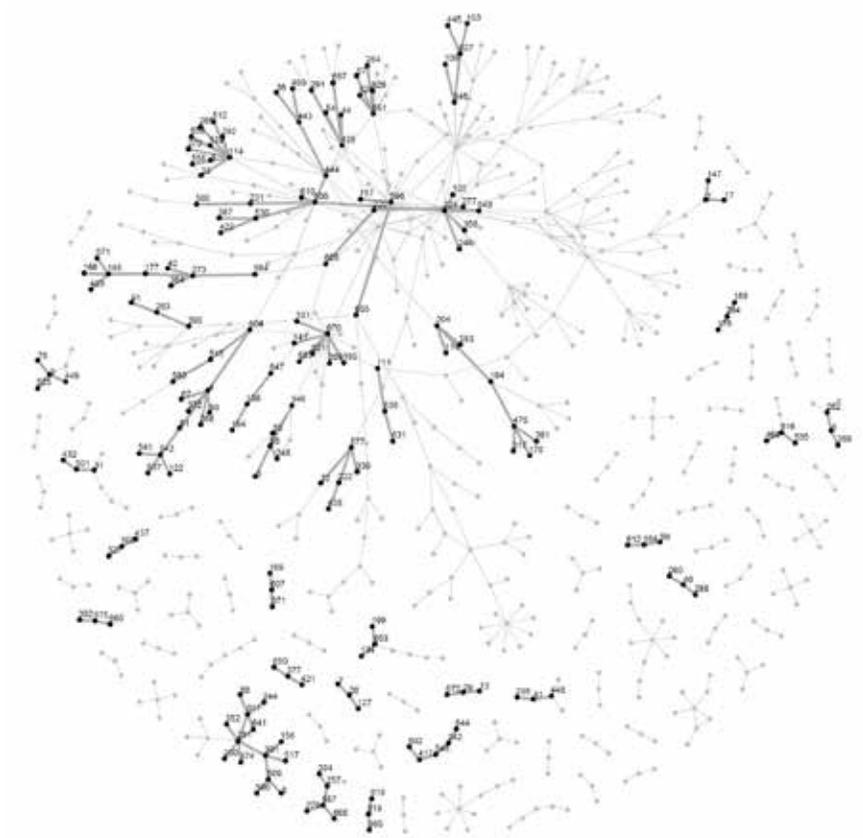


Figura 4 Grafo representativo dos acadêmicos envolvidos nas bancas de defesa de trabalhos do grupo 2, com foco na discussão de implicações das avaliações

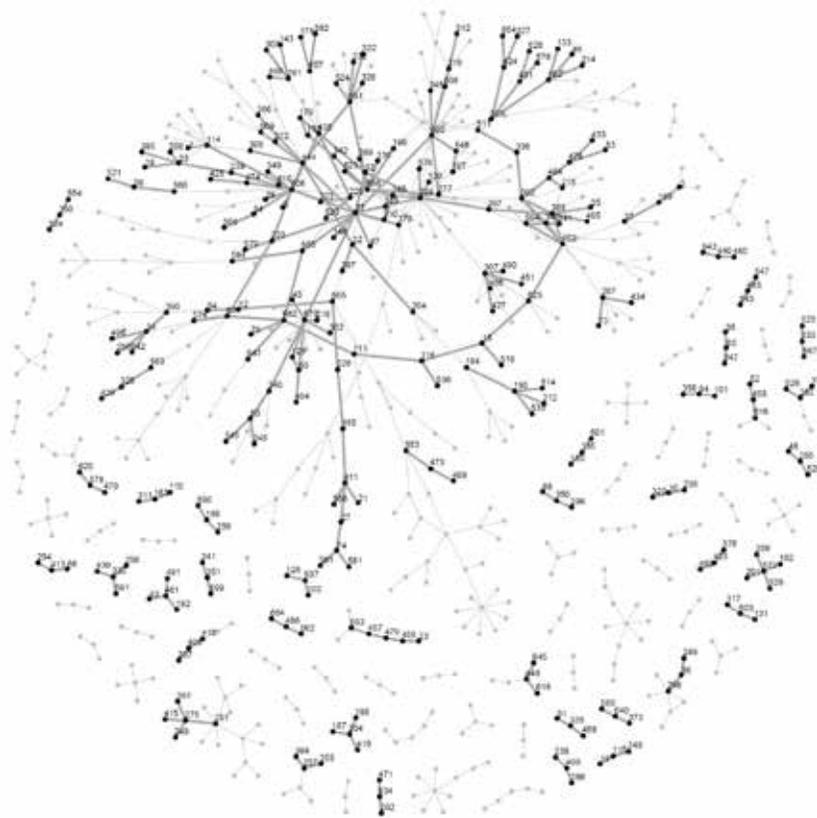


Figura 5 Grafo representativo dos acadêmicos envolvidos nas bancas de defesa de trabalhos do grupo 3, com foco na discussão de aspectos metodológicos das avaliações

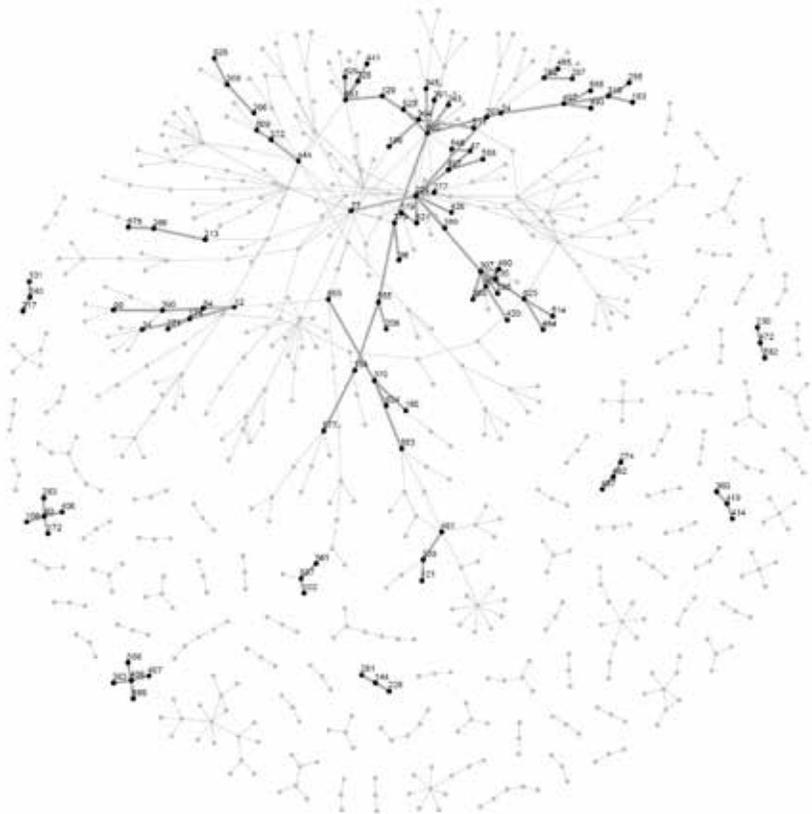
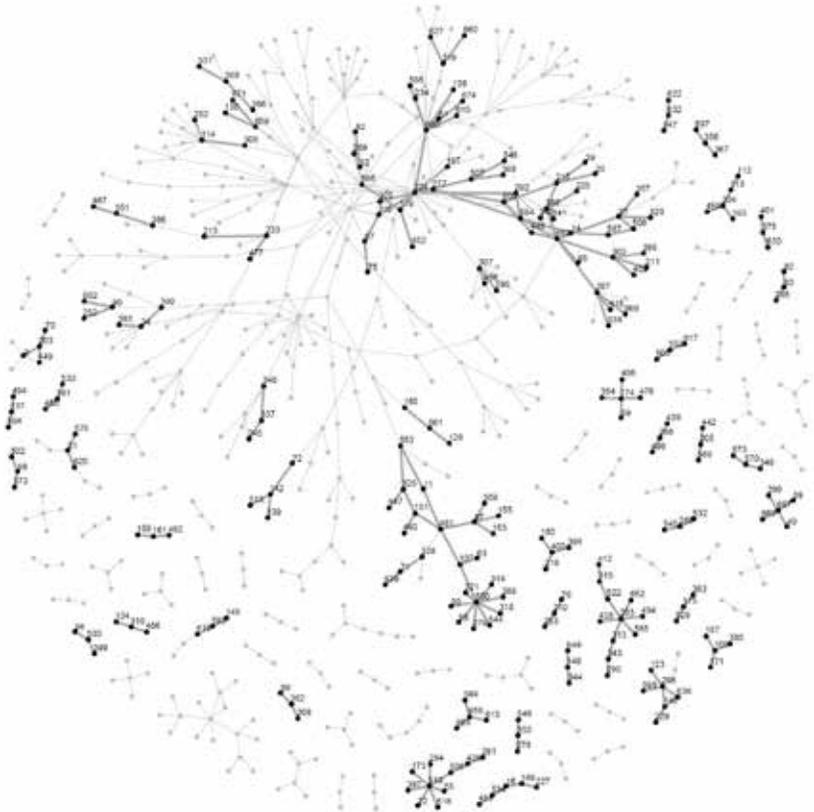


Figura 6 Grafo representativo dos acadêmicos envolvidos nas bancas de defesa de trabalhos do grupo 4, com foco em outros objetos de pesquisa



Fonte: base de dados da pesquisa. Elaboração: Eduardo Capocchi. Organização da autora.

A tabela 7 sumariza a participação relativa das instituições de ensino e pesquisa mais representativas em cada grupo, possibilitando analisar se há instituições que têm uma produção mais consolidada em cada um dos subgrupos. Segundo os dados obtidos, observa-se que a PUCSP aparece como a instituição que mais produziu nos grupos 1 e 4. A USP destaca-se no grupo 2 e a UnB nos trabalhos do grupo 3.

Tabela 7 Participação absoluta das instituições por subgrupo temático

	GRUPO 1	TRABALHOS	GRUPO 2	TRABALHOS	
46% dos trabalhos	PUCSP	6	USP	11	43% dos trabalhos
	UNESP/Araraquara	5	UFMG	7	
	UNICAMP	5	PUC Rio	7	
	UFMG	4	PUCSP	6	
	UFJF	3	UnB	4	
	UnB	3	UNESP/Marília	4	
	Outras 26 instituições	31	USP/Ribeirão Preto	4	
			UFJF	4	
			Outras 44 instituições	62	
	Total	57	Total		
	GRUPO 3	TRABALHOS	GRUPO 4	TRABALHOS	
54% dos trabalhos	UnB	6	PUCSP	13	36% dos trabalhos
	UFJF	5	UFPE	6	
	PUC Rio	4	PUC Rio	5	
	UEM	3	UFMG	4	
	UFPE	3	UFJF	4	
	Outras 15 instituições	18	Outras 33 instituições	57	
	Total	39	Total	89	

Fonte: base de dados da pesquisa.

Conclusão

Pela natureza interdisciplinar da temática, observa-se o aumento expressivo da produção de teses e dissertações sobre avaliação externa e em larga escala. No entanto, a análise dos grafos produzidos pela abordagem da análise de redes sociais mostra que há uma tendência de dispersão dessa produção, não se podendo afirmar que ela se origina apenas em grupos de pesquisa e em instituições com maior tradição de investigação na área. Tal apropriação da temática por representantes de diferentes áreas e linhas de pesquisa, ao mesmo tempo em que pode propiciar que esse

objeto de estudo seja analisado à luz de uma diversidade de metodologias, referenciais teóricos e formas de constituição dos problemas de pesquisa, contribuindo para ampliação do campo e aprofundamento das discussões a ele inerentes, pode também levar à piora da qualidade dos trabalhos na medida em que a produção isolada pode apresentar limites na análise e interpretação dos problemas e questões inerentes a esse tipo de avaliação.

Observa-se, também, que a produção acadêmica parece estar limitada às lentes teórico-metodológicas usuais para alguns subgrupos. Enquanto se observa que o subgrupo 3 tende a produzir conhecimento por meio da utilização de metodologias quantitativas e mistas, a produção dos demais subgrupos temáticos é, essencialmente, realizada por meio de abordagens qualitativas.

Finalmente, observa-se que, apesar de alguns referenciais teóricos serem reconhecidos como referências tradicionais e obrigatórias ao estudo da temática, a baixa frequência de citações desses autores já “clássicos” permite afirmar que diversos trabalhos têm sido produzidos sem que se dialogue, de alguma forma, com os autores de referência na discussão das avaliações externas e em larga escala. Esse aspecto pode levar ao enfraquecimento teórico do campo, hipótese que seria corroborada pelo alto número de citações de obras isoladas, quer produzidas por autores individualmente, quer por meio de colaboração entre diversos autores.

Tal situação vem corroborar o que Alves (1992, p.54) já afirmara há uma década: a revisão bibliográfica é um dos aspectos mais frágeis na análise das teses e dissertações. Segundo a autora, “[a revisão da literatura] tem por objetivo iluminar o caminho a ser trilhado pelo pesquisador, desde a definição do problema até a interpretação dos resultados”.

Não seria de se esperar que a produção acadêmica, no nível de mestrado e de doutorado, apresentasse a visão mais completa possível do estado de conhecimento, das tensões, das contradições e das confluências de pensamento já consolidadas acerca de uma temática?

Se assim fosse, seria importante que o pesquisador conhecesse a produção relacionada ao seu objeto de estudos, dialogasse com ela, analisando criticamente o estado do conhecimento já produzido, procurando identificar os pontos de consenso, as controvérsias, as lacunas, os aspectos

que precisam ser esclarecidos, ou seja, os problemas de pesquisa que realmente contribuem para o desenvolvimento do conhecimento na área (LAROCA; ROSSO; SOUZA; 2005; ALVES, 1992). E, para garantir isso, tanto o orientador quanto a banca têm papel fundamental.

Considera-se, assim, primordial o papel da comunidade científica, representada pelo orientador e pela banca examinadora dos trabalhos, para a orientação dos estudantes de mestrado e de doutorado no que se refere ao diálogo com os conhecimentos já produzidos no campo, bem como na garantia da adequação do trabalho aos padrões esperados em relação a uma dissertação de mestrado ou tese de doutorado. Como lembra Alves (1992), orientadores e membros de banca assumem papel primordial na garantia da qualidade dos produtos produzidos no âmbito do mestrado e do doutorado e, conseqüentemente, na validação de conhecimentos que possam ser incorporados ao campo de estudo.

No entanto, o que se observa, neste estudo, é que a maioria dos trabalhos têm sido orientados e avaliados em comunidades isoladas de pesquisadores, muitos dos quais com pouca tradição na pesquisa da temática, o que não garante que as análises garantam o nível de qualidade esperado para os trabalhos acadêmicos produzidos no âmbito da pós-graduação. Caberia, então, que os orientadores atentassem para a indicação dos membros da banca, a fim de garantir que haja maior interlocução com o conhecimento já produzido o que propiciaria, em tese, melhor qualificação dos trabalhos.

Referências

ANDRÉ, Marli. E. D. Pesquisa e educação: buscando rigor e qualidade. *Cadernos de Pesquisa*, n. 113, p. 51-64, jul. 2001.

ALVES, Alda J. A “revisão da bibliografia” em teses e dissertações: meus tipos inesquecíveis. *Cadernos de Pesquisa*, São Paulo, n. 81, p. 54, maio 1992.

ALVES-MAZOTTI, Alda J. Relevância e aplicabilidade da pesquisa em educação. *Cadernos de Pesquisa*, n. 113, p. 39-50, jul. 2001.

BAUER, Adriana. Estudos sobre Sistemas de Avaliação Educacional no Brasil: um retrato em preto e branco. *Revista @mbienteeducação*, São Paulo, v. 5, p. 7-31, 2012.

_____. Avaliação de redes de ensino e gestão educacional: o que apontam os estudos acadêmicos. IV CONGRESSO IBERO-AMERICANO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO; In: VII CONGRESSO LUSO BRASILEIRO DE POLÍTICA E ADMINISTRAÇÃO DA EDUCAÇÃO, *Anais...*Lisboa, 2014.

BAUER, Adriana; REIS, Adriana T. Balanço da produção teórica sobre avaliação de sistemas educacionais no Brasil: 1988 a 2011. In: 36ª REUNIÃO NACIONAL DA ANPED, *Anais...* 2013, Goiânia/GO.

BOURDIEU, Pierre. *Los usos sociales de la ciencia: por una sociología clínica del campo científico*. Buenos Aires: Ediciones Nueva Visión, 2003.

BORDIN, Andréa S.; GONÇALVES, Alexandre L.; TODESCO, José L. Análise da colaboração científica departamental através de redes de coautoria. *Perspectivas em Ciência da Informação*, v. 19, n. 2, p. 37-52, abr./jun. 2014.

BORGATTI, Stephen. P.; EVERETT, Martin G.; JOHNSON, Jeffrey C. *Analyzing social networks*. Los Angeles, Thousand Oaks, Calif.; London: SAGE Publications, 2013.

CONWAY, Drew. *Social Network Analysis in R*. 2009. Disponível em: <http://www.databaser.net/moniwiki/pds/DataAnalysis/social_network_analysis_02.pdf>. Acesso em: 16 ago. 2014.

CSARDI, Gábor; NEPUSZ, Tamás. The igraph software package for complex network research. *InterJournal*, v. Complex Systems, p. 1695, 2006. Disponível em: <<https://pdfs.semanticscholar.org/1d27/44b83519657f5f2610698a8ddd177ced4f5c.pdf>>. Acesso em: 1º abr. 2018.

DIAS SOBRINHO, José. Campos e caminhos da avaliação: a avaliação da educação superior no Brasil. IN: FREITAS, L. C. de (Org.). *Avaliação: construindo o campo e a crítica*. Florianópolis: Insular, 2002.

FIALHO, Joaquim M. R. Análise de redes sociais: princípios, linguagem e estratégias de ação na gestão do conhecimento. *Perspectivas em Gestão & Conhecimento*, João Pessoa, v. 4, Número Especial, p. 9-26, out. 2014.

GATTI, Bernardete A. Implicações e perspectivas da pesquisa educacional no Brasil contemporâneo. *Cadernos de Pesquisa*, n. 113, p. 65- 81, jul. 2001.

HAAS, Peter. *Introduction: epistemic communities and international policy coordination*. International Organization, vol. 46, no. 1, *Knowledge, Power, and International Policy Coordination*. (Winter), p. 1-35, 1992.

LAROCCA, Priscila; ROSSO, Ademir J.; SOUZA, Audrey P. A formulação dos objetivos de pesquisa na pós-graduação em Educação: uma discussão necessária. *Revista Brasileira de Pós-Graduação*, Brasília: Capes, v. 2, n. 3, p. 118-133, mar. 2005.

MARQUES, Eduardo. Redes Sociais e instituições na construção do estado e da sua permeabilidade. *Revista Brasileira de Ciências Sociais*, v. 14, n. 41, 1999, p. 45-67.

MARTELETO, Regina M. Análise de redes sociais: aplicação nos estudos de transferência da informação. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 71-81, 2001.

MARTINS, Dalton L. *Análise de redes sociais de colaboração científica no ambiente de uma federação de bibliotecas digitais*. 2012. Tese (Doutorado em Cultura e Informação)–Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2012.

MOLINA, José L.; et. al. *Talleres de autoformación con programas informáticos de análisis de redes sociales*. UAB: Barcelona, 2005. Disponível em: <http://revista-redes.rediris.es/webredes/red_tematica/talleresars.pdf>. Acesso em: 1º abr. 2018.

PENNA NETO, Manoel C. de O.; FREY, Klaus; CZAJKOWSKI JR., Sérgio. Avaliação estrutural das redes sociotécnicas. In: Tamara Tania Cohen. (Org.). *Ciberpólis: redes no governo da cidade*. Rio de Janeiro: 7 Letras, v. 1, p. 47-69, 2007.

RIGHETTI, Sabine. Brasil cresce em produção científica, mas índice de qualidade cai. *Folha de S. Paulo*, 22 abr. 2013. Disponível em: <http://tools.folha.com.br/print?site=emcimadahora&url=http://www1.folha.uol.com.br/ciencia/2013/04/1266521-brasil-cresce-em-producao-cientifica-mas-indice-de-qualidade-cai.shtml>. Acesso em: 27 maio 2014.

SACERDOTE, Helena C. de S. *A mediação segundo Feuerstein e o uso da informação em educação on-line*. 2018. Tese (Doutorado em Ciência da Informação)–Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2018.

SANTOS, Nálbia de A.; FARIAS, Manoel R. S. *Modelos meta-teóricos para estudos epistemológicos do processo de pesquisa acadêmica*. In: X CONGRESSO USP DE CONTROLADORIA E CONTABILIDADE. São Paulo: Fipepapi, 26 e 27, jul. 2010.

SCRIVEN, Michel. The Concept of a Transdiscipline: And of Evaluation as a Transdiscipline. *Journal of MultiDisciplinary Evaluation*, v. 5, n. 10, p. 65-66, July 2008.

SILVA, Maximiliano B. *Redes sociais de colaboração acadêmica: tendências, determinantes e fricções do processo de formação de coautorias entre pesquisadores no Brasil*. 2013. Tese (Doutorado em Teoria Econômica)–Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2013.

SOCZKA, Luís. Modelos de análise de redes sociais e limitações do modelo de equilíbrio estrutural de Heider. *Psicologia*, Lisboa, v. 15, n. 1, p. 147-175, jan. 2001.

SOUZA, Queila R.; QUANDT, Carlos O. Metodologia de análise de redes sociais. In: DUARTE, Fábio; QUANDT, Carlos; SOUZA, Queila (Org.). *O Tempo das Redes*. São Paulo: Perspectiva, p. 31-63, 2008.

ANEXOS

Tabela 8 Dados descritivos de cada sub-rede

	Nº DE PARTICIPANTES NA SUB-REDE	Nº DE PARTICIPANTES NA REDE MAIOR	% DE PARTICIPANTES DA REDE MAIOR EM RELAÇÃO AO TOTAL	Nº DE SUB-REDES ISOLADAS*	TOTAL DE PARTICIPANTES NAS SUB-REDES ISOLADAS	% DE PARTICIPANTES DAS REDES ISOLADAS	Nº DE SUB-REDES INTERMEDIÁRIAS**
GRUPO 1	179	22	12,3	27	91	50,84	6
GRUPO 2	274	114	41,6	46	149	54,38	1
GRUPO 3	103	53	51,5	13	43	41,75	1
GRUPO 4	241	50	20,7	43	142	58,92	4

Fonte: base de dados da pesquisa.

Nota: * sub-redes isoladas contemplam de 3 a 5 participantes

** sub-redes intermediárias contemplam de 6 a 21 participantes

Quadro 1 Correspondência número-índice e nomes dos pesquisadores (orientadores e/ou membros de banca)

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
1	Abramowicz_M	33	Arelaro_LRG	67	Betlinski_C
2	Abrucio_FL	34	Arena_DB	68	Bezerra_AA
3	Adrião_TMd	35	Arenhardt_LO	69	Bezerra_MA
4	Aguilar_LE	36	Arraes_RdA	70	Bianchetti_L
5	Alavarse_OM	37	Arretche_MTd	71	Bianchini_BL
6	Albuquerque_HMd	38	Arruda_EEd	72	Bicalho_DC
7	Albuquerque_MGM	39	Assunção_JJ	73	Bittencourt_MVL
		40	Avelar_LMd	74	Bittencourt_NA
8	Almeida_AMF	41	Avelino_Fo_G	75	Boing_LA
9	Almeida_MdL	42	Azevedo_JL	76	Bolzan_DPV
10	Almeida_PCd	43	Azevedo_TMd	77	Bonamino_AMC
11	Almouloud_SA	44	Baptista_MC	78	Boneti_LW
12	Alvarenga_GM	45	Baquero_RVA	79	Borges_LFF
13	Alvarenga_MSd	46	Barbosa_MCD	80	Borges_ON
14	Alves_FCd	47	Barbosa_MEN	81	Borges_ZP
15	Alves_JMP	48	Barbosa_RLL	82	Bortoni_ME
16	Alves_ML	49	Barreira_C	83	Bortoni-Ricardo_SM
17	Alves_MA	50	Barreiro_IMd		
18	Alves_NG	51	Barretto_ESd	84	Botler_AMH
19	Amâncio_LNd	52	Barros_HFd	85	Boubleur_JP
20	Amaral_CAF	53	Barros_RPd	86	Braga_FM
21	Ambrozetti_NB	54	Barroso_J	87	Brandão_RHL
22	Amorim_Fo_OB	55	Barroso_LC	88	Brasil_WB
23	Amui_LBL	56	Baruffi_AMZ	89	Brayner_FHA
24	Andrade_DFd	57	Bastos_AVB	90	Britto_LPL
25	Andrade_EdC	58	Bastos_HFB	91	Brocanelli_CR
26	Andrade_FRB	59	Bastos_JB	92	Brooke_NPd
27	André_MED	60	Bastos_MHC	93	Bryan_NAP
28	Aquino_JRG	61	Bastos_RR	94	Bueno_BAd
29	Arana_HG	62	Beato_Fo_CC	95	Bueno_MSS
30	Araújo_AJd	63	Bellemain_PMB	96	Burgos_MB
31	Araújo_DLd	64	Bellini_LM	97	Buriasco_RLC
32	Araújo_GC	65	Belluzzo_Jr_W	98	Bussad_WdO
		66	Betini_GA	99	Cabral_ALT

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
100	Cabral_FA	136	Cicillini_GA	171	Dias_EF
101	Cabral_LS	137	Cintra_AMM	172	Dias_MdG
102	Caldas_AdR	138	Coelho_Fo_CAd	173	Diniz_AMA
103	Calderano_MdA	139	Corrêa_HT	174	Dionísio_AP
104	Calderón_AI	140	Corrêa_MV	175	Dobranszky_EA
105	Calmon_PCD	141	Corsetti_B	176	Domene_AC
106	Campos_RHd	142	Coscarelli_CV	177	Dourado_LF
107	Campos_TMM	143	Cóssio_MdF	178	Duarte_AMC
108	Candau_VMF	144	Costa_HG	179	Duarte_MRT
109	Canen_A	145	Costa_Md	180	Dutra_IM
110	Canôas_JW	146	Costa_MdA	181	Dynnikov_CMS
111	Capanema_CdF	147	Costa_VMF	182	Elliot_LG
112	Capovilla_AGS	148	Coutinho_CdQ	183	Espinheira_PL
113	Capovilla_FC	149	Cruz_MNd	184	Esteban_MT
114	Cappelletti_IF	150	Cunha_STd	185	Evangelista_O
115	Cardoso_CJ	151	Curi_E	186	Eyng_AM
116	Carnielli_BL	152	Cury_CRJ	187	Fachin_RC
117	Caro_AMP	153	Cury_HN	188	Fancul_MCd
118	Carvalho_JM	154	Cyranka_LFd	189	Farah_MFS
119	Carvalho_JBP	155	Cyrino_MCd	190	Farenzena_N
120	Carvalho_JCB	156	Dalben_AML	191	Farias_MdS
121	Carvalho_MTd	157	Dalben_ÂIL	192	Fávero_AA
122	Castanho_SEM	158	Dalcin_A	193	Feitosa_MD
123	Castelar_Lld	159	DallÁglio-Hatt- nher_MM	194	Feltran_RCd
124	Castro_MFC			195	Fernandes_CdO
125	Castro_MLS	160	D´Ambrosio_U	196	Fernandes_CMB
126	Castro_MRd	161	Darsie_MM	197	Fernandes_CAC
127	Catrib_AMF	162	David_A	198	Fernandes_DC
128	Cavazotti_MA	163	David_CM	199	Fernandes_DMB
129	César_CC	164	Davies_N	200	Fernandes_FS
130	Cesar_MRd	165	Davis_CLF	201	Fernandes_CdO
131	Chaves_EOd	166	Deitos_RA	202	Fernandes_R
132	Chizzotti_A	167	Del_Roio_MT	203	Fernandes_SRd
133	Chrispino_A	168	Demartini_ZdB	204	Ferraço_CE
134	Cianflone_ARL	169	Demo_P	205	Ferreira_Jr_A
135	Ciasca_MIF	170	Dias_AA	206	Ferreira_ES

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
207	Ferreira_FdH	243	Gameran_D	279	Guimarães_GL
208	Ferreira_LA	244	Gandini_RPC	280	Günther_H
209	Ferreira_MOV	245	Ganzelli_P	281	Gutierrez_RH
210	Ferreira_VS	246	Garchet_MB	282	Haas_CM
211	Ferro_AR	247	Garcia_CB	283	Haase_VG
212	Fetzner_AR	248	Garcia_TMF	284	Hadad_RM
213	Fiad_RS	249	Garita_RMS	285	Hage_SAM
214	Firme_TP	250	Gebran_RA	286	Heinig_OLd
215	Firpo_SP	251	Gentilini_JA	287	Hildenbrand_LMA
216	Fischer_BTD	252	Geraldi_CMG	288	Hofling_EdM
217	Flôres_OC	253	Gesser_V	289	Holanda_MC
218	Florissi_S	254	Ghedin_EL	290	Horikawa_AY
219	Fogaça_A	255	Giampaoli_V	291	Hypolito_ÁLM
220	Fonseca_M	256	Giglio_CMB	292	Iannone_LR
221	Foster_MMd	257	Giovinazzo_Jr_CA	293	Ikeda_SN
222	Frade_ICA	258	Giroto_CGG	294	Inforsato_EdC
223	França_RLd	259	Gisi_ML	295	Jacobi_PR
224	Franco_Jr_FCJ	260	Giubilei_S	296	Jeffrey_DC
225	Franco_AMd	261	Givisiez_GHN	297	Jesus_DMd
226	Franco_MLP	262	Godoy_MGG	298	Jorge_Nt_PdM
227	Freitas_AGB	263	Gomes_AA	299	Junqueira_L
228	Freitas_ALP	264	Gomes_CAd	300	Kassouf_AL
229	Freitas_DGF	265	Gomes_CMA	301	Kerbaui_MTM
230	Freitas_Dd	266	Gomes_IZ	302	Klein_Remi
231	Freitas_DNT	267	Gonçalves_FdO	303	Klein_Ruben
232	Freitas_KSd	268	Gonçalves_JF	304	Klering_LR
233	Freitas_LCd	269	Gonçalves_LAO	305	Kramer_S
234	Freitas_MTd	270	Gonçalves_MDd	306	Krawczyk_NR
235	Freitas_VAd	271	Gorni_DAP	307	Laros_JA
236	Frota_FHd	272	Gouveia_AB	308	Lasmar_C
237	Fürkotter_M	273	Gracindo_RV	309	Leite_MCL
238	Furlanetto_MM	274	Grando_NI	310	Lelis_IAO
239	Fusari_JC	275	Grego_SMD	311	Lemes_SdS
240	Gabriel_R	276	Gregolin_MdR	312	Lemgruber_MS
241	Gama_RP	277	Gremaud_AP	313	Lessa_ABC
242	Gama_ZJ	278	Guedes_MdC	314	Lessa_MML

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
315	Liberali_FC	351	Manrique_AL	387	Meneguello_R
316	Lier-De_Vitto_MFd	352	Manzano_E	388	Menezes_Fo_NA
317	Lima_Jr_LPd	353	Marchezan_RMF	389	Menezes_JE
318	Lima_APd	354	Marcuschi_E	390	Menin_AMM
319	Lima_CRO	355	Marin_AJ	391	Micarello_HAL
320	Lima_HMF	356	Marinho_M	392	Michelotto_RM
321	Lima_PF	357	Marques_EC	393	Miller_S
322	Lima_RAP	358	Marques_LR	394	Minhoto_MAP
323	Limongi_FMP	359	Marques_MAd	395	Minto_CA
324	Longhin_Thomazi	360	Marques_MOd	396	Miranda_SHG
325	Lopes_CAE	361	Marques_W	397	Miranda-Ribeiro_A
326	Lopes_DA	362	Marquesi_SC	398	Monastério_LM
327	Lopes_MLM	363	Marquezzine_MC	399	Monteiro_HM
328	Lordelo_JAC	364	Marteleteo_LF	400	Monteiro_SMM
329	Lourencetti_GdC	365	Martins_AS	401	Monteiro_SB
330	Lourenço_FA	366	Martins_ÂM	402	Moraes_MAF
331	Luce_MBM	367	Martins_AA	403	Moraes_MCM
332	Lugli_RSG	368	Martins_MAR	404	Moraes_ScD
333	Luna_SVd	369	Martins_MAV	405	Morais_AGd
334	Macedo_ECd	370	Massi_CDB	406	Moreira_AF
335	Macedo_MdS	371	Masson_MAC	407	Moreira_LC
336	Machado_AF	372	Matos_BTP	408	Morelatti_MRM
337	Machado_LM	373	Mattos_JMd	409	Mori_NNR
338	Machado_SDA	374	Mattos_MJV	410	Morosini_MC
339	Maciel_FIP	375	Maximo_AC	411	Moroz_M
340	Madalozzo_RC	376	Mcdonald_BC	412	Motta_LMV
341	Madeira_RdA	377	Melo_AAS	413	Mourão_ARB
342	Mafra_LdA	378	Melo_KLR	414	Mutim_ALB
343	Magalhães_MCC	379	Melo_MFP	415	Muzzeti_LR
344	Magina_SMP	380	Melo_MPd	416	Nacarato_AM
345	Magrone_E	381	Melo_MMd	417	Najjar_JNV
346	Maia_GZA	382	Melo_SDG	418	Nakabashi_L
347	Maia_GBd	383	Mendes_JR	419	Nascimento_AD
348	Maia_MLA	384	Mendonça_EF	420	Nascimento_Ed
349	Mainardes_J	385	Mendonça_SGd	421	Nascimento_EFV
350	Malavasi_MMS	386	Menegassi_RJ		

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
422	Nascimento_MEP	458	Passador_JL	494	Ramos_RdC
423	Nascimento_RBd	459	Paula_CSd	495	Rangel_MdA
424	Nasser_L	460	Paulo_RM	496	Raphael_HS
425	Negreiros_PJL	461	Paviani_NMS	497	Raposo_MCF
426	Neves_MAC	462	Pazello_ET	498	Rasia_JM
427	Neves_MMB	463	Pena_LPM	499	Rauen_FJ
428	Ninin_MOG	464	Pereira_Jr_A	500	Rego_LML
429	Noronha_OM	465	Pereira_CdC	501	Reinaldo_MAG
430	Novaes_LC	466	Pereira_LMd	502	Reinhard_N
431	Novelino_MSf	467	Pereira_RF	503	Reyes_CR
432	Nunes_MLd	468	Perez_MdL	504	Rezende_WM
433	Oliveira_AMH	469	Peroni_VMV	505	Riani_JdL
434	Oliveira_BCd	470	Pessoa_ACR	506	Ribeiro_Fo_CAC
435	Oliveira_DA	471	Petratti_P	507	Ribeiro_LCd
436	Oliveira_ELd	472	Pierson_AHC	508	Ribeiro_MdG
437	Oliveira_LG	473	Pietropaolo_RC	509	Ribeiro_R
438	Oliveira_MAd	474	Piletti_N	510	Rigotti_JIR
439	Oliveira_MCM	475	Pimenta_SdA	511	Rios_Nt_ELG
440	Oliveira_MCA	476	Pinto_APP	512	Rios_MPG
441	Oliveira_NF	477	Pinto_ALG	513	Rocha_GAS
442	Oliveira_Rd	478	Pinto_CCd	514	Rocha_MPd
443	Oliveira_RTC	479	Pinto_JMd	515	Rocha_MSP
444	Oliveira_RLP	480	Pinto_NB	516	Rodrigues_Jr_JF
445	Oliveira_RdF	481	Pires_CMC	517	Rodrigues_MdM
446	Orellano_VIF	482	Pirola_NA	518	Rodrigues_RdN
447	Ortigão_MIR	483	Pizzi_LCV	519	Rodrigues_RL
448	Pacheco_RSV	484	Ponczek_V	520	Rodrigues_SBP
449	Paim_MMW	485	Pontes_MGd	521	Rodriguez_MV
450	Paiva_MAV	486	Pontal_LLF	522	Romero_TRd
451	Palacios_KEP	487	Portanova_R	523	Ronca_ACC
452	Palis_GdL	488	Portes_ÉA	524	Rosa_DL
453	Palma_Fo_JC	489	Prado_MEB	525	Rosa_GAd
454	Pantaleoni_NTd	490	Rabelo_ML	526	Rosito_MMB
455	Parisotto_ALV	491	Ramos_FB	527	Rothen_JC
456	Pasquali_L	492	Ramos_MG	528	Ruiz_AR
457	Passador_CS	493	Ramos_MN	529	Ruiz_EMS

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
530	Rus_Perez_JR	566	Serbino_RV	602	Sobrinho_JD
531	Russo_MH	567	Setubal_MA	603	Soffner_RK
532	Sabino_CdV	568	Sforni_MSd	604	Sordi_MRL
533	Sad_LA	569	Shimazaki_EM	605	Sousa_JVd
534	Saes_DAM	570	Shimoishi_JM	606	Sousa_SZL
535	Sales_JBd	571	Shiroma_EO	607	Souza_AME
536	Salinas_JL	572	Sicca_NAL	608	Souza_APF
537	Salles_FC	573	Sicsú_AL	609	Souza_ÂRd
538	Sampaio_AV	574	Silva_Jr_CAd	610	Souza_CPd
539	Sampaio_CDd	575	Silva_AFd	611	Souza_CBG
540	Sampaio_JdR	576	Silva_ACd	612	Souza_DBD
541	Sandano_W	577	Silva_CSR	613	Souza_LES
542	Sanfelice_JL	578	Silva_DBD	614	Souza_MIG
543	Santana_MdM	579	Silva_Dd	615	Souza_RC
544	Santiago_MML	580	Silva_ERd	616	Souza_RJd
545	Santos_Fo_JCd	581	Silva_EFd	617	Spinillo_AG
546	Santos_Jr_O	582	Silva_HCd	618	Spink_PK
547	Santos_IMd	583	Silva_IMd	619	Stauffer_AdB
548	Santos_JAd	584	Silva_MAd	620	Suassuna_L
549	Santos_JBC	585	Silva_MAC	621	Sztajn_P
550	Santos_MCd	586	Silva_Md	622	Tafner_PSB
551	Santos_MdC	587	Silva_NdV	623	Talim_SL
552	Santos_SL	588	Silva_PLd	624	Taurino_MdS
553	Santos_VdM	589	Silva_PBG	625	Tavares_HR
554	Sargentini_VMO	590	Silva_SR	626	Tavares_MR
555	Sarmento_DC	591	Silva_VBd	627	Teixeira_BdB
556	Sarti_FM	592	Silva_WCd	628	Teixeira_MdP
557	Sass_O	593	Simões_RHS	629	Tenório_RM
558	SauL_AMA	594	Siqueira_JHS	630	Tibali_EFA
559	Savioli_AMP	595	Soares_Nt_JJ	631	Tollini_MI
560	Scaff_EAd	596	Soares_JF	632	Tolosa_HC
561	Scheibe_L	597	Soares_MB	633	Toschi_MS
562	Schneider_L	598	Soares_RB	634	Trindade_LMS
563	Schnetzler_RP	599	Soares_SLF	635	Tróccoli_BT
564	Scorzafave_LGD	600	Soares_TM	636	Trois_SdC
565	Scotton_MT	601	Sobral_AEB	637	Trojan_RM

ÍNDICE	NOME	ÍNDICE	NOME	ÍNDICE	NOME
638	Trompieri_Nt_N	651	Verhine_RE	664	Vitória_MIC
639	Utsumi_MC	652	Vian_Jr_O	665	Wachowicz_LA
640	Uyeno_EY	653	Viana_ABN	666	Waltemberg_FD
641	Vaidergorn_J	654	Vianna_CCd	667	Wanderley_JCV
642	Valente_SMP	655	Vianna_HM	668	Warde_MJ
643	Valente_WR	656	Victor_EdF	669	Weber_S
644	Valle_BdB	657	Vieira_LMF	670	Werle_FOC
645	Vaz_JC	658	Vieira_MdT	671	Xavier_Fa_C
646	Veiga_Fo_AdL	659	Vieira_MCT	672	Xavier_MdC
647	Veiga_IPA	660	Vilarinho_LRG	673	Yamashita_Y
648	Veiga_Ld	661	Villardi_RM	674	Yazbeck_DCd
649	Vendramini_CMM	662	Villas_Boas_BMd	675	Zainko_MAS
650	Verçosa_EdG	663	Viola_SEA	676	Zanini_M

Adriana Bauer

Doutora em Educação pela Faculdade de Educação da Universidade de São Paulo, com estágio pós-doutoral no Institute of Education (IoE), Londres. Pesquisadora da Fundação Carlos Chagas e Docente na Universidade de São Paulo, São Paulo (SP), Brasil;
dri_bauer@yahoo.com.br

Artigo submetido em 26/04/2018

Aprovado em 01/06/2018

INTEGRAÇÃO LEITURA/ESCRITA NA PROVA DE REDAÇÃO DO VESTIBULAR UNICAMP: AMEAÇAS AO CONSTRUTO?

READING/WRITING INTEGRATION IN THE UNICAMP ENTRANCE
EXAMINATION: THREATS TO THE CONSTRUCT?

INTEGRACIÓN LECTURA/ESCRITURA EN LA PRUEBA DE REDACCIÓN DEL
EXAMEN DE INGRESO DE UNICAMP: AMENAZAS AL CONSTRUTO?

Matilde Virginia Ricardi Scaramucci

Luciana Amgarten Quitzau

RESUMO

Considerando que diferentes propostas de produção escrita variam de uma edição para outra de um mesmo exame e que essas variações podem ameaçar o seu construto, este artigo apresenta resultados de uma pesquisa que analisa tarefas integradas de leitura e escrita da prova de redação do Vestibular Unicamp, ou mais especificamente, os enunciados das tarefas e a natureza dos textos-insumo que as acompanha. As duas tarefas selecionadas – uma potencialmente adequada e outra potencialmente problemática – foram analisadas em conjunto com as grades de correção e com os textos produzidos por examinandos a propósito dessas tarefas. Concluímos que variações distintas não só ameaçam o construto do exame como também comprometem as representações dos examinandos sobre as tarefas, afetando seu desempenho, o processo de correção e os resultados da avaliação.

Palavras-chave: avaliação; tarefas integradas; redação; ameaça ao construto; vestibular.

ABSTRACT

Considering that different writing tasks vary from one edition to the other in the same exam and that these variations could pose threats to the construct assessed, this article presents the results of a research that analyzes integrated tasks of reading and writing in the Unicamp entrance examination. More specifically, we analyzed the way the tasks were constructed and the nature of the reading texts that come along with them. The two selected tasks – one considered potentially adequate and the other potentially problematic – were analyzed together with the rubrics and the texts produced by the test takers. We concluded that the distinct variations not only threaten the construct of the exam but also affect the test takers' representation of the tasks, their performance, the rating process and the results of assessment.

Keywords: assessment; integrated tasks; essay; threats to the construct; university entrance examination.

RESUMEN

Considerando que las diferentes propuestas de escritura varían de una edición para otra de un mismo examen y que estas variaciones pueden amenazar el constructo de ese examen, este artículo expone los resultados de una investigación que analiza tareas de escritura y lectura en el examen de redacción del examen de ingreso de Unicamp. Más específicamente, analizamos los enunciados de las tareas y la naturaleza de los textos fuentes. Las dos tareas seleccionadas – una potencialmente adecuada y otra potencialmente problemática – han sido analizadas de forma conjunta con las grillas de evaluación utilizadas y los textos producidos por los examinandos. Concluimos que las variaciones distintas amenazan el constructo del examen como también comprometen las representaciones de los examinandos sobre las tareas, afectando su performance, la corrección y los resultados de la evaluación.

Palabras clave: evaluación; tareas integradas; redacción; amenazas al constructo; examen de ingreso.

Introdução

A avaliação da produção escrita em exames de desempenho – exames que avaliam as habilidades de maneira direta – envolve grande complexidade e, por isso mesmo, tem desafiado avaliadores, principalmente quando essa produção pressupõe também a leitura de textos, como tem ocorrido em alguns exames nacionais de larga escala, tais como a prova de redação do Exame Nacional do Ensino Médio (Enem) e exames vestibulares (Universidade Estadual de Campinas – Unicamp, por exemplo), no contexto de língua materna; e em exames de proficiência em língua estrangeira, como, por exemplo, o exame Celpe-Bras (Certificado de Proficiência em Língua Portuguesa para Estrangeiros). No caso específico da prova de redação do Vestibular Unicamp, exame focalizado neste artigo, o texto que acompanha o enunciado de cada tarefa de escrita é denominado texto-insumo (ou texto-fonte) e seu papel está claramente explicitado. Em outras palavras, a prova de redação da Unicamp é claramente uma prova de leitura e escrita como habilidades integradas.

Enquanto muitos recursos têm sido investidos no processo de correção, com treinamento de corretores e elaboração de instrumentos ou grades de avaliação, muito pouco tem sido feito, em contexto nacional, no que diz respeito à análise da adequação das propostas, envolvendo os enunciados e a natureza dos textos que os acompanham (PILEGGI, 2015; SCARAMUCCI, 2016).

A avaliação de habilidades integradas tem sido uma das tendências da avaliação contemporânea, na medida em que, na vida real, as habilidades geralmente são utilizadas de maneira integrada. O uso de tarefas integradas pode contribuir para o aumento da autenticidade e da validade do exame. Segundo Plakans (2008) e Weigle, Yang e Montee (2013), as tarefas que integram leitura e escrita, além de fornecerem conteúdo para o texto dos examinandos, são mais parecidas com as atribuições da escrita acadêmica, visto que geralmente a escrita, na academia, é feita a partir da leitura de outros textos. Weigle, Yang e Montee (2013) apontam também outra vantagem das tarefas integradas: elas fornecem múltiplos meios de o examinando demonstrar proficiência linguística e habilidades. Plakans (2008) salienta, contudo, que embora mais exames estejam adotando as tarefas integradas de leitura e escrita, há algumas dificuldades a serem enfrentadas,

como a escolha dos textos-insumo e o seu uso literal, ao que adicionaríamos a necessidade de maior definição do construto.

Vários autores buscaram definir o que são essas tarefas (ASENCIÓN DELANEY, 2008; CUMMING, 2013; GEBRIL; PLAKANS, 2013). Nesta pesquisa, assumimos a definição de Knoch e Sitajalabhorn (2013, p. 306), que propõem um sentido mais específico para o conceito ao constatarem que muitas tarefas diferentes são chamadas de “integradas”. Segundo os autores:

Tarefas de escrita integrada são tarefas nas quais os examinandos são apresentados a um ou mais textos-insumo ricos em língua, e são levados a produzir composições escritas que requerem (1) garimpagem dos textos-insumo, (2) seleção de ideias, (3) síntese de ideias de um ou mais textos-insumo, (4) transformação da língua usada no insumo, (5) organização de ideias e (6) uso de convenções estilísticas tais como conexão de ideias e reconhecimento de fontes. A grade usada para dar nota a tais composições precisa levar em conta esses aspectos específicos a tarefas de escrita integrada (tradução do autor).¹

Diferentes propostas de produção escrita podem apresentar configurações distintas dessas habilidades, o que pode estar em conflito com o construto e ser responsável por resultados divergentes mesmo para examinandos de mesma habilidade. Além disso, os textos incluídos nas propostas podem ser de naturezas distintas, introduzindo variáveis que podem causar discrepâncias nos resultados. O construto a ser avaliado (produção escrita e leitura, neste caso) é operacionalizado em itens e propostas e também em suas grades e procedimentos de correção. O que se deseja é que o construto avaliado seja mantido de maneira consistente nas várias etapas do exame, o que nem sempre ocorre (SCARAMUCCI, 2017).

Enunciados que não mantêm, de forma consistente, o construto, ou que introduzem variações de uma prova para outra podem comprometer os

¹ *Integrated writing tasks are tasks in which test takers are presented with one or more language-rich source texts and are required to produce written compositions that require (1) mining the source texts for ideas, (2) selecting ideas, (3) synthesizing ideas from one or more source texts, (4) transforming the language used in the input, (5) organizing ideas and (6) using stylistic conventions such as connecting ideas and acknowledging sources. The rating scale used to grade such compositions needs to take account of these features specific to integrated writing tasks.*

resultados, assim como as inferências feitas com base nesses resultados. Essas variações – causadas por ambiguidades, falta de explicitação, idiossincrasias, entre outros – poderão introduzir o que a literatura tem denominado de “ameaças à validade” dos construtos avaliados (MESSICK, 1989; SCARAMUCCI, 2011; SCARAMUCCI, 2017). Essas variações podem levar examinandos a interpretações (e, portanto, representações) distintas das demandas da proposta, que podem levar a desempenhos também distintos, comprometendo os resultados e as inferências a serem feitas com base nesses resultados, mesmo quando esses examinandos têm as mesmas habilidades de leitura e escrita, neste caso; ou a resultados semelhantes, quando, na realidade, os examinandos têm habilidades distintas. As variações nas propostas podem ter impacto negativo também no processo de correção, dificultando o ajuste das grades, comprometendo sua qualidade e isenção (SCARAMUCCI, 2017).

Embora esteja fora do escopo desta proposta, vale a pena mencionar que as variações nas propostas podem comprometer não apenas a qualidade da avaliação, mas também os efeitos que esta pode ter nas práticas de ensino que antecedem os exames, potencializando os negativos e minimizando os positivos (o que a literatura tem denominado de validade consequencial, da qual os efeitos retroativos são um exemplo).

Neste artigo, estruturado em quatro partes, relatamos os resultados de um projeto de pesquisa cujo objetivo geral foi investigar em que medida propostas de produção escrita (incluindo o enunciado e os textos-insumo), em diferentes edições do exame de redação da Unicamp, podem estar avaliando diferentes configurações das habilidades de leitura e escrita, levando examinandos com diferentes habilidades a serem classificados em um mesmo nível de proficiência.

As perguntas de pesquisa que nortearam o projeto são:

A prova de redação da Unicamp apresenta variações ou divergências em suas propostas (enunciados e textos), de uma edição para outra? De que natureza são essas variações?

As diferentes edições desse exame podem estar avaliando “habilidades diferentes” no que está sendo considerado “mesmo nível de proficiência”? Essas variações ameaçam o construto do exame?

A Unicamp foi uma das universidades pioneiras não apenas na introdução de uma redação em seu exame vestibular, nos anos 80, mas também na maneira como essa prova foi concebida, incluindo uma coletânea de excertos de textos para leitura para fomentar a elaboração da dissertação, e algumas vezes, até mesmo, por textos completos para subsidiar o texto narrativo e a carta-argumentativa. Entre três propostas, o examinando escolhia qual gostaria de desenvolver. Esse modelo permaneceu inalterado até o vestibular de 2004, quando o exame passou a fornecer uma coletânea única para as três propostas. Em 2010, o exame adquiriu o formato que mantém atualmente, passando a exigir não mais a escolha de uma proposta, mas a produção obrigatória de três gêneros discursivos distintos². A cada gênero corresponderia uma tarefa diferente a ser cumprida, que integraria leitura e escrita.

O documento interno com a proposta de mudança para o atual modelo, encaminhado à Câmara Deliberativa do Vestibular (SCARAMUCCI; LOPES; FRUNGILLO, 2010), elencava como justificativas: (1) a dificuldade em fazer três propostas (dissertação, carta argumentativa e narrativa) que tivessem o mesmo nível de dificuldade; (2) a dificuldade em avaliar a leitura pela coletânea, uma vez que grande parte das redações produzidas apresentavam como evidência de leitura apenas citações e informações recuperadas, o que não permitiria avaliar se o candidato tinha realmente capacidade de ler textos inteiros que lhe seriam exigidos na universidade; (3) o subaproveitamento dos excertos e o uso de fórmulas para lidar com eles; (4) a mescla de “tipos textuais” (dissertação e narrativa) e “gêneros discursivos” (carta-argumentativa)³; (5) a constatação de que a permanência do mesmo modelo de prova por 20 anos estava trazendo um efeito retroativo negativo para o ensino: a criação de “receitas” de textos no lugar do desenvolvimento de práticas de leitura e escrita.

Ao apresentar a proposta de reformulação da sua prova de Redação, o Vestibular Unicamp destacou, de maneira mais incisiva do que em 1987, o seu objetivo de avaliar leitura e escrita de forma integrada, através da leitura e produção de gêneros discursivos diversos:

2 Passaram a ser dois gêneros discursivos no Vestibular de 2013.

3 O documento registra que, no caso da dissertação, essa indefinição seria ainda maior, pois haveria os que a consideram um “tipo textual” e os que já definem dissertação escolar como um novo gênero.

A leitura continuaria a ser avaliada juntamente com a escrita, porém não mais com base em uma coletânea geral, e sim em textos também de gêneros diversos, de forma que a compreensão do texto seja imprescindível para a elaboração da tarefa. [...] Um aspecto importante, nesse caso, é fazer com que o gênero do texto base para a leitura seja sempre diferente do gênero do texto a ser produzido, de forma a desestabilizar a prática corrente de “cortar e colar” trechos do texto sem reflexão (SCARAMUCCI; LOPES; FRUNJILLO, 2010).

As principais características da prova atual de Redação do Vestibular Unicamp são, portanto: (1) a solicitação de produção de gêneros distintos; (2) a associação de cada um desses gêneros a uma tarefa distinta; (3) a presença, no enunciado da tarefa, de informações a respeito de seu propósito, dos interlocutores envolvidos e do gênero a ser produzido; e (4) o fato de cada tarefa ser acompanhada por um ou mais textos-insumo cujos propósitos de leitura também estão expressos no enunciado.

Nessa prova, os parâmetros de correção são as próprias condições de produção: o gênero solicitado e a interlocução que o caracteriza, o propósito da tarefa, a leitura e a articulação escrita, com atenção especial para a modalidade e a coesão. De acordo com Scaramucci, Lopes e Frunjillo (2010, p. 7 e 8), em relação à leitura, o candidato deveria estabelecer pontos de contato dos textos fornecidos com o texto produzido por ele, deixando clara a relevância desses pontos para o seu projeto de texto. Não seria permitida a mera reprodução total ou parcial desses textos, sem que houvesse um trabalho de elaboração sobre eles. Em relação ao propósito, ele deveria demonstrar que compreendeu as instruções que recortam o tema e a motivação da escrita. Quanto ao gênero e à interlocução, seria avaliado se o projeto de texto do candidato estaria de acordo com as características do gênero solicitado e dos interlocutores implicados nele. Em relação à escrita, seriam avaliados a adequação do conjunto lexical e o uso de estruturas sintático-semânticas com a finalidade de construir um texto cuja leitura fosse fluida e envolvente.

Na avaliação do cumprimento das condições de produção, temos uma avaliação binária: cumpriu/não cumpriu (1/0). O zero, nos critérios Propósito,

Interlocução e Gênero, implica a anulação do texto, ou seja, ele não passará pela avaliação de qualidade, não recebendo uma nota na avaliação holística. A anulação total da prova de redação do candidato só acontece se ele tiver os dois textos anulados.

Na avaliação da qualidade do texto produzido, leva-se em conta, fundamentalmente, a autoria e a modalidade escrita/coesão, que são compreendidas da seguinte forma (COMVEST, 2010):

Autoria: domínio da prática de leitura e escrita que permite explorar produtivamente, por meio de um projeto de texto, as condições de produção dadas (propósito, interlocução e gênero), conferindo autonomia ao texto.

Modalidade/Coesão: aspectos formais (elementos coesivos e adequação do conjunto lexical ao padrão normativo da língua e ao gênero solicitado) que garantem a unidade textual, resultando em uma fluidez da leitura que não é afetada por eventuais impropriedades.

É importante esclarecer que diferentes configurações desses elementos determinam os 6 níveis/descriptores da grade que avalia a qualidade da formulação escrita (0 ponto – muito ruim – a 5 pontos – muito bom).

As duas etapas da avaliação – de cumprimento dos propósitos de escrita e da qualidade desse cumprimento – foram materializadas em duas grades distintas: a específica e a holística. O corretor primeiramente avalia o texto por meio da grade específica e, em seguida, por meio da grade holística.

As provas de redação do Vestibular Unicamp são corrigidas de maneira independente por dois corretores que não têm acesso a qualquer informação sobre o autor do texto, o curso no qual pretende ingressar e a avaliação feita pelo 2º corretor. Cada correção vale 8 pontos e, em caso de divergência maior que 2 pontos, o texto passa por uma 3ª correção, feita pelos professores responsáveis pelo treinamento dos corretores. Nesse caso, a nota final do candidato será uma composição da nota do 3º corretor com a que for mais próxima da dele.

Método

A metodologia do estudo foi essencialmente qualitativa. O fato de termos quantificado o número de textos para análise, como mostraremos mais adiante, não significa que nosso estudo seja quantitativo, pois não utilizou técnicas quantitativas de análise.

Após leitura e revisão preliminar da literatura sobre habilidades integradas e análise de documentos referentes ao exame da Unicamp (manuais do candidato, enunciados de provas, expectativas das bancas, documentos internos e grades de correção), elaboramos um protocolo de análise com os seguintes aspectos: propósito, enunciador, interlocutor, gênero, suporte, conhecimento prévio, texto-insumo, autenticidade, escrita e leitura. Nessa análise preliminar, como nosso foco de trabalho era o estudo do uso do texto-insumo para o cumprimento dos propósitos de escrita (avaliação da escrita integrada à leitura), não nos detivemos em aspectos mais formais, como modalidade e coesão. Em um estudo futuro, talvez seja interessante observar como as características da escrita do(s) texto(s)-insumo influenciam a escrita dos examinandos.

Em seguida, analisamos todas as propostas aplicadas a partir do Vestibular 2011, num total de 14 enunciados diferentes. A análise desses enunciados nos levou a estabelecer hipóteses a respeito de qual(is) tarefa(s) promoveria(m) melhor integração das habilidades de leitura e escrita e qual(is) poderia(m) comprometer o construto da prova, seja por falta de clareza, pelo tópico tratado, pelas características dos enunciados das propostas e dos textos-insumo (CHO; RIJMEN; NOVAK, 2013), seja por apresentar ambiguidades, falta de explicitação e idiosincrasias capazes de ameaçar a validade desse construto (SCARAMUCCI, 2011). Essa análise também envolveu a natureza dos textos-insumo que acompanham os enunciados. Pudemos constatar que há textos de gêneros diversos, como gráfico, entrevista, crônica, matérias jornalísticas. Dada essa variedade, a extensão dos textos também varia, assim como os contextos em que originalmente circulam. Portanto, outro critério de escolha das tarefas foi que tivessem textos-insumo de origem distinta (matérias jornalísticas e fábula), em quantidade diferente (uma era acompanhada de dois textos-insumo e a outra por um), e que exigissem a mesma habilidade: a síntese. Os textos-insumo, nesse caso, são de dificuldade semelhante. Se a dificuldade desses textos fosse distinta nas propostas, seria mais um elemento complicador para nossa comparação.

A análise prévia nos permitiu selecionar uma tarefa considerada potencialmente adequada (Texto 2 do Vestibular 2013) e outra potencialmente inadequada (Texto 1 do Vestibular 2016), ambas reproduzidas nos Anexos I e II. Em seguida, solicitamos à Comvest a liberação do *corpus*, composto por 3 mil textos, sendo 1.500 redações de cada uma das duas propostas selecionadas, com suas respectivas notas finais (assim como as notas parciais que cada um dos corretores deu) e as grades de correção usadas nessas duas edições do exame.

Assim que tivemos acesso ao *corpus* cedido pela Comvest, que era bastante extenso, fizemos uma filtragem inicial a partir da planilha de notas que nos foi cedida pela instituição. Selecionamos 50 textos de cada proposta, usando como critério de busca aqueles cuja nota de Propósito tinha variado entre 0 e 1 nas duas primeiras correções. Acreditamos que a discrepância de notas nesse critério poderia sinalizar textos que apresentaram alguma dificuldade de correção, seja por características dos enunciados e das grades, seja por dificuldades dos próprios candidatos. Propósito é um critério importante, pois é ele que determina quais serão os objetivos de leitura do(s) texto(s)-insumo. Como já dissemos, o propósito da tarefa é explicitado ao examinando no enunciado da prova. Sendo assim, como queremos observar habilidades de leitura e escrita, esse nos parece ser o critério mais adequado para ser observado. Como consequência, os textos analisados podem ter ficado com nota final considerada abaixo da média (1, 2 ou 3) ou mediana (4, 5 ou 6), caso uma 3ª correção tenha confirmado a nota 1. Para entender o que estava sendo efetivamente avaliado como adequado às expectativas, demonstrando o que seria integrar habilidades nesse exame, analisamos também todas as redações com nota máxima (8). Como elas aparecem em quantidade muito pequena na amostra, recorreremos aos livros elaborados pela banca examinadora com as melhores redações dos vestibulares 2013 e 2016⁴, através dos quais tivemos acesso a 20 textos considerados exemplares (10 de cada proposta).

Como mencionado anteriormente, a grade específica se limita a atribuir 0 ou 1 ponto aos critérios Propósito, Interlocução e Gênero. Desta forma, ela faz uma “filtragem” dos textos, avaliando se o examinando conseguiu trabalhar com o que a banca elaboradora considera o mínimo para cada

⁴ São eles: *Vestibular Unicamp: Redações 2013 e Vestibular Unicamp: Redações 2016*, publicados pela Comvest em parceria com a Pró-Reitoria de Graduação da Unicamp.

um desses aspectos, e que está estabelecido nos descritores da grade específica. Em nossa análise, buscamos avaliar se examinandos e corretores enfrentariam alguma dificuldade em propostas com problemas no enunciado ou na natureza do texto-insumo selecionado (textos com poucas informações, ou muito metafóricos, ou com linguagem prolixa, etc.). No caso dos corretores, nosso objetivo era verificar se havia obstáculos para estabelecer parâmetros de correção.

Resultados

Apresentamos, a seguir, cada uma das propostas selecionadas para entender em que medida havia variações entre uma e outra e de que natureza eram essas variações. Em seguida, analisamos os textos produzidos por examinandos em resposta a cada uma das tarefas para observar se houve impacto dos enunciados e dos textos-insumo no seu desempenho.

O Texto 2, do Vestibular 2013, apresenta um enunciado claro, com potencial para ser considerado um bom instrumento de avaliação da leitura e da escrita de forma integrada. Em nossa análise, a proposta conta com os seguintes pontos positivos:

- O gênero cuja produção é solicitada (carta do leitor) é bastante diferente do gênero dos textos-insumo (matéria jornalística), o que dificulta o aparecimento de redações que se limitem a “copiar e colar” trechos do insumo;
- A solicitação da recuperação de informações de forma contextualizada (“fazer menção à matéria publicada, de modo que quem não a tenha lido entenda a importância da crítica que você faz”) mostra para o examinando que a recuperação de informações também é uma habilidade que se espera dele, mas que ela deve ser feita quando faz sentido para o desenvolvimento do gênero;
- O uso de um 2º texto para criticar o 1º, levando o candidato a comparar conteúdos, perceber implícitos, argumentar, dentre outras habilidades exigidas.

Dadas essas características que consideramos positivas, pensamos que, quando houvesse divergência entre os corretores em relação ao cumprimento do Propósito, isso aconteceria por uma desatenção pontual e com pouca frequência, e seria facilmente identificado pelo 3º corretor.

Na proposta em questão, de acordo com a grade específica, para ter nota 1 em Propósito, o texto deveria cumprir os seguintes requisitos:

- Especificar a matéria publicada no jornal.
- Apresentar crítica a essa matéria.
- Fundamentar essa crítica nos dados da 2ª matéria.
- Relacionar a crítica ao consumo de álcool por adolescentes.

Temos, a seguir, o exemplo de um texto acima da média, em que o examinando conseguiu cumprir esses requisitos com bastante clareza:

Exemplo 1

Meu nome é J.S. e sou um leitor assíduo da Folha de São Paulo. Como tal, me senti incomodado com a matéria “Cães vão tomar uma ‘gelada’ com cerveja pet” (1), pois nela não há nenhuma alusão aos possíveis efeitos que esse tipo de produto pode ter sobre o consumo de álcool, sobretudo entre os adolescentes (2).

Venho acompanhando o debate sobre o álcool na adolescência, e toda essa minha preocupação vem, em grande parte, dos resultados de uma pesquisa, que foram publicados recentemente na matéria “Vergonha Nacional”, da revista Veja. Eles mostram que um em cada três jovens brasileiros de 14 a 17 anos se embabedou ao menos uma vez no período de um ano, situação que, a meu ver, é alarmante (3).

O lançamento de uma cerveja específica para cães, aliado à forma com que a notícia foi apresentada, ressalta o aspecto positivo do consumo de bebidas alcólicas. Vide a primeira sentença da matéria: “Nada é melhor

que uma cervejinha depois de um dia de cão". Esse destaque dado à função social do consumo de álcool induz ao uso desta como justificativa para tal ato, e não apenas por adolescentes, mas também por adultos que presenciavam jovens se embriagando. A mesma pesquisa que citei também mostra que em 17% dos episódios em que menores se embebedaram estavam acompanhados dos próprios pais ou tios (3).⁵

Compreender implícitos é uma atividade mais complexa do que localizar informações pontuais; sendo assim, uma das habilidades exigidas nos textos acima da média era perceber que, como cachorros fazem parte da família, se eles podem beber, será difícil negar uma cerveja ao adolescente, fato que pode contribuir para a piora do quadro exposto pelo Texto 2. Essa proposta permitiu que o corretor tivesse clareza ao diferenciar textos acima da média, que enxergam esses implícitos, daqueles que se restringem a habilidades que exigem menor trabalho do leitor, como a recuperação de informações. A seguir, apresentaremos um trecho em que o examinando claramente demonstra ter a habilidade de compreender implícitos:

Exemplo 2

Fiquei incomodado, pois a matéria se preocupou apenas com a informação sobre o novo produto, exótico e desnecessário, em minha opinião. Faltou ao redator a preocupação de também analisar as consequências sociais, como a possibilidade de aumento do consumo de álcool na sociedade, principalmente por parte dos adolescentes. Se adultos, donos de animais de estimação, consentem em dar cerveja aos cães para ter companhia em momentos de descontração, eles também podem permitir o consumo de álcool a filhos ou sobrinhos menores de idade (grifos nossos).

Nos exemplos a seguir veremos trechos que levaram os corretores a dar notas diferentes no critério Propósito. No primeiro, um dos corretores não considerou que o texto critica o produto cerveja para cães e não a matéria publicada pelo jornal *Folha de São Paulo*. Esse problema, identificado no cruzamento das correções, pode ser facilmente resolvido no processo de 3ª correção.

5 Grifos nossos. Eles destacam: (1) a especificação da matéria publicada no jornal, (2) a crítica a essa matéria, sendo que essa crítica está relacionada ao consumo de álcool por adolescentes, e (3) a fundamentação da crítica com dados da 2ª matéria.

Exemplo 3

Ao ler a matéria “Cães vão tomar uma ‘gelada’ com cerveja pet” resolvi enviar-lhes esta carta como um acompanhante do debate sobre o álcool na adolescência, sendo contra esse produto.

Exemplo 4

Adolescentes menores de idade acham normal o consumo de álcool, e provavelmente começam a usufruir dessas bebidas por causa da convivência e influência de parentes e amigos.

Nesse trecho, provavelmente, os corretores ficaram em dúvida a respeito da presença ou não de dados do 2º texto para fundamentar a crítica. É importante observar que, embora o trecho claramente tenha sido produzido inspirado no conteúdo da matéria da revista *Veja*, ele ignora os dados trazidos por ela. Nesse caso, um corretor deve ter considerado como um dado o trecho “começam a usufruir dessas bebidas por causa da convivência e influência de parentes e amigos”, enquanto o outro não considerou. Trata-se, então, de uma redação limite, o que explica a divergência de notas. É importante destacar que o enunciado deixava claro o trabalho que deveria ser feito, mas o examinando não mostrou habilidade de relacionar os dois textos.

Como exemplo de proposta potencialmente inadequada, selecionamos o Texto 1, do Vestibular 2016. Alguns aspectos chamaram a nossa atenção durante a análise desse enunciado:

- O detalhamento excessivo sobre o concurso de resenhas, que pode confundir o candidato: (1) Por que é necessário informar que a resenha fará parte de um concurso? (2) E que esse concurso é promovido pelo Centro de Apoio ao Estudante? (3) E que a função desse órgão é desenvolver atividades culturais na Faculdade? (4) E que o objetivo do concurso é estimular a leitura de obras literárias e ampliar o horizonte cultural dos estudantes? (5) E que uma comissão julgadora escolherá dez textos?.

- O uso de negritos que destacavam aspectos secundários do enunciado.
- A falta de autenticidade da tarefa, pois dificilmente veremos circular a resenha de uma fábula, pois esse gênero não apresenta, geralmente, produtos culturais tão curtos.
- A exigência de construção de uma analogia que, tal como foi solicitada, talvez também não possa ser avaliada por critérios objetivos. Como avaliar quais são os pontos de contato do projeto de um candidato com a fábula apresentada, visto que o texto fornecido é essencialmente metafórico e abre o leque de interpretações possíveis?

Como registrado anteriormente, optamos por observar o Propósito, pois o trabalho com o texto-insumo está mais direcionado a esse aspecto. Na prova da Unicamp, o enunciado direciona a leitura que o candidato deve fazer do(s) texto(s)-insumo, logo, não é possível cumprir o propósito de escrita sem a leitura desses textos; é o enunciado que determina qual é o propósito de leitura do(s) texto(s)-insumo e que tipo de trabalho deve ser feito com ele(s). No entanto, é importante registrar outro elemento que dificultou a avaliação dessa proposta: o critério Gênero, no Vestibular Unicamp, não avalia o uso do texto-insumo, mas, na proposta em questão, há a exigência de produção de uma síntese. Embora o texto-insumo seja utilizado para a elaboração da síntese, o trabalho com ele não é avaliado pela grade. Além disso, apesar de muitos trabalhos no exterior mencionarem que sínteses e resumos são solicitados como integração de leitura e escrita (CUMMING, 2013; YU, 2013; ASENSIÓN DELANEY, 2008; SPIVEY; KING, 1989), não há consenso sobre o seu uso em avaliações em larga escala. Yu (2003) afirma que o resumo é um processo fundamental envolvido na realização de qualquer tarefa de avaliação integrada, mas é de natureza multidimensional e requer diferentes parâmetros para avaliar a sua qualidade. Além dessa questão, também observamos que a síntese não se caracteriza como um gênero, mas como uma habilidade necessária para a produção de diversos gêneros, por isso talvez a sua presença não deva ser o único elemento avaliado para verificar se um texto se configura como resenha ou não.

Em relação ao critério Propósito, são avaliados dois aspectos, no caso dessa proposta:

- A construção da situação social análoga à da fábula (que envolva um problema coletivo);
- A elaboração de um fechamento para o texto que tenha relação com a temática da fábula.

Essas exigências podem ter imposto alguns desafios para a correção, tais como a dificuldade para avaliar se o problema é coletivo e análogo ao da fábula. Observamos que a quantidade de divergência na nota de Propósito é maior na amostra desse texto no que na do anterior (apresentaremos os números na seção 4) e percebemos que a avaliação dessa proposta parece ter dependido mais do julgamento do corretor do que da aplicação da grade, pois cabe a ele dizer se o problema é coletivo e se a situação é análoga. Fazer isso sem que se estabeleçam pontos de contato com o texto-insumo pode ser bastante difícil.

Apresentamos a seguir exemplos de divergência na avaliação do Propósito:

Exemplo 5

Situações análogas à da fábula ocorrem no Brasil atual quando o assunto é política. Muitos discutem os problemas e possuem soluções simples e eficazes, mas quando ações diferentes de falar são necessárias, todos se calam. Há grandes assembleias⁶ e excelentes soluções, mas não existem aqueles dispostos a realmente colocá-las em prática. A política brasileira entre as pessoas comuns é uma grande reunião de ratos de La Fontaine.

Exemplo 6

Problemáticas semelhantes a essa não são raras e nem recentes na sociedade. A organização de protestos significativos no Brasil, por exemplo, é desproporcional às numerosas queixas diárias da população. E, com o desenvolvimento das tecnologias de informação e o advento das redes

6 Os trechos são transcrições fiéis dos textos originais.

sociais, a facilidade com que as pessoas podem organizar medidas práticas aumentou consideravelmente, contudo não há a iniciativa necessária.

Exemplo 7

Essa fábula é uma crítica à sociedade e que compreende a geração atual. Enxergamos o problema na comunidade e muitas vezes a solução, porém o individualismo não os permite reagir.

Todos os exemplos mostram dificuldades na avaliação dessa proposta: como definir se a questão exposta é uma situação análoga? Tratar de problemas genéricos, como corrupção na política e violência, é algo comum na amostra, no entanto não se pode dizer que basta mencionar esses aspectos mais gerais para que uma situação específica tenha sido criada.

No exemplo a seguir, os corretores consideraram a situação exposta como análoga, mas tiveram dificuldade para qualificá-la, visto que um corretor atribuiu nota 2 (ruim) pela grade holística e outro, nota 5 (muito bom).

Exemplo 8

Uma situação social análoga e atual é a do assédio sexual que ocorre em diversos ambientes de trabalho em que uma pessoa de cargo mais alto se insinua para outra de maneira direta ou indireta, pois o/a abusado(a) sabe que se não ceder, isto pode culminar no seu desemprego.

Todos os exemplos anteriores são de textos da amostra que nos foi disponibilizada pela Comvest, enquanto que os próximos (5 e 6) foram extraídos do livro que divulga as melhores redações do vestibular, já mencionado anteriormente. Em comum, os dois últimos têm o fato de estabelecerem claramente pontos de contato com elementos significativos da fábula que tornariam a situação análoga, como o medo, a opressão, a desigualdade de forças, a necessidade de tomar atitude, a reunião, a falta de alguém que tenha coragem para ser o primeiro a reagir. Quanto mais explícito for esse trabalho, melhor será a analogia. Talvez tenha faltado, no enunciado, a explicitação de que era necessário estabelecer pontos de contato com a fábula na construção dessa analogia.

Exemplo 9

Na fábula de La Fontaine, “A deliberação tomada pelos ratos”, os ratos vivem sob o terror do gato Rodilardo, chegando a se assemelhar ao terror vivido na França durante o governo de Robespierre. Rodilardo, assumindo uma política do medo, matava os ratos que se atreviam a sair dos buracos das paredes onde se escondiam. Os ratos, assim como diversos povos (além dos franceses) que viviam sob o controle de um Estado autoritário, opressor e violento, sentindo-se insatisfeitos e encurralados discutem uma possível solução para a situação deplorável em que vivem.

Apesar de acreditarem ter chegado num meio de melhorar significativamente suas vidas, os ratos, que são também o povo oprimido, voltam a um impasse, ninguém se dispõe a lutar e sofrer por um bem maior.

Exemplo 10

O grande número de casos de abuso de violência policial é resultado de um mecanismo semelhante ao ilustrado na fábula, pois as pessoas, aterrorizadas por servidores públicos que deveriam protegê-las, isolam-se em suas casas, de modo muito semelhante aos ratos em suas tocas, perdendo oportunidades e passeios porque têm medo de sair. Sem sair de casa, muitos utilizam a Internet para divulgar os abusos da polícia. Estes são, na maioria dos casos, aconselhados a denunciar o crime, mas eles, por temer serem agredidos pela autoridade para a qual denunciaram, não seguem tal conselho, permitindo que os que abusaram da força, concedida a eles pelo Estado, violentem outras pessoas.

Na análise dos textos da amostra, também observamos casos de fechamentos da resenha que foram comprometidos pelo enunciado. Como houve um investimento grande em explicar a finalidade do concurso de resenhas, alguns examinandos podem ter pensado que essa questão era muito importante e que seria avaliada. Sendo assim, a usaram no fechamento do texto, muitas vezes forçando uma relação entre a fábula e a importância da leitura. Veremos isso nos exemplos a seguir, extraídos de textos com notas holísticas que variaram entre 2 (ruim) e 3 (razoável):

Exemplo 11

A literatura é uma grande ponte para que possamos ter mais consciência sobre a sociedade que vivemos. Mas ela vai além disso, é preciso entrar no universo da literatura para descobrir tudo o que ela abriga. Há livros clássicos, como os de Machado de Assis, mas também contos de natal, gibiz, histórias de ficção científica e muito mais. Porém não basta só colocá-los na estante, é preciso ler, se aventurar na cultura.

Exemplo 12

A leitura é fundamental para a reflexão e adquirir conhecimento, especialmente sobre o mundo. Explorar os problemas sociais e discuti-los é algo que a literatura faz muito bem e ela quem nos transporta para um mundo que pode ser mais real do que o que estamos.

Exemplo 13

Nesse contexto, é imprescindível que os estudantes conheçam a importância da leitura, não só na vida acadêmica, mas em suas vidas pessoais, uma vez que ela é uma forma de ampliação de horizontes culturais. Por tudo isso, a fábula de La Fontaine, assim como muitas outras, adquire importância por estar intimamente relacionada a problemas cotidianos e por fornecer análises e explicações sobre diversos tipos de comportamentos comumente encontrados. Dessa forma, sua leitura, assim como a de outras obras literárias, se torna uma importante ferramenta de evolução pessoal e coletiva.

A falta de qualidade dos fechamentos acima fica mais explícita se comparados com um exemplo de bom fechamento:

Exemplo 14

Quando deparados com o medo, os ratos, que buscavam soluções como leões, voltam ao seu tamanho inicial e à sua insignificância, e os homens, ao desistirem da luta, também voltam a ser ratos, em estado de menoridade e acomodados com uma situação abominável, porque, enfim dá menos trabalho.

Exemplo 15

La Fontaine, com esta fábula, transmite a moral de que, embora seja importante deliberar os assuntos, é imprescindível executá-los. Situação semelhante ocorre quando uma comunidade enfrenta problemas com a segurança pública. Em um determinado bairro com alto índice de violência, pouco adianta lastimar-se dos crimes ocorridos ou discutir soluções em uma rede social. Caso este alto índice de violência ocorra em razão da ausência de escolas ou atividades culturais, essa comunidade deverá se organizar e levar os fatos às autoridades competentes para que providenciem o necessário e, com a participação de todos, seja resolvido concretamente o problema.

Discussão

Na tarefa de 2013, há duas matérias, uma de jornal e outra de revista. Para criticar uma é necessário utilizar dados da outra. Na de 2016, há um único texto, uma fábula. Ao resenhá-la, o examinando deve construir uma situação social análoga. Ao contrário da outra proposta, em que havia elementos de dois textos para avaliar a leitura, nesta tarefa, o candidato deve entender a metáfora e construir algo além dela. O que pode ser um exercício interessante para a sala de aula torna-se um elemento complicador para a avaliação exigida em um exame de larga escala. A proposta de 2013 exigiu um uso mais complexo dos textos-insumo, trabalhando com implícitos. Examinandos mais proficientes conseguiram realizar essa tarefa, mas os menos proficientes não, pois se limitaram a recuperar informações – o que é muito bom, pois a tarefa conseguiu discriminar os níveis de proficiência. Já a proposta de 2016 exigiu, além da síntese (que foi avaliada no critério Gênero e não no Propósito), uma situação análoga. A síntese foi ignorada, e é na sua elaboração que poderíamos observar se, de fato, ocorreu a integração das habilidades. Nesse caso, o enunciado da proposta pôs grande ênfase em conhecimentos externos (a situação análoga), tirando o foco da leitura, o que não ocorreu na proposta de 2013. Essas variações no enunciado geraram grande diferença nas produções e problemas para a correção. Como a grade holística da Unicamp vincula as notas altas ao uso de texto-insumo para o cumprimento do Propósito, não encontramos nas notas altas textos que tenham negligenciado ou interpretado de forma insatisfatória o(s) texto(s)-insumo.

Levando em consideração o que foi exposto, é interessante observar a diferença na quantidade de textos cujas avaliações divergiram no critério Propósito (apresentaram notas 0 e 1) nas amostras de textos produzidos em resposta às propostas selecionadas de 2013 e 2016. Na proposta de 2013, considerada boa, apenas 56 textos de um total de 1.500 apresentaram divergência de nota no critério Propósito. Já na proposta de 2016, considerada problemática, 259 apresentaram esse problema. Não se pode afirmar que uma mudança tão significativa se deva à variação de corretores na banca. Acreditamos que ela se deva à dificuldade apontada nesta pesquisa de se definir algumas questões essenciais para a avaliação desse critério, que são: (1) a situação retratada é social? (2) é coletiva? (3) é análoga à da fábula? (4) como lidar com fechamentos “forçados”, pois o examinando se sentiu obrigado a usar os negritos do texto? Com exceção da última questão, todas as outras envolvem subjetividade na correção.

As redações de 2016 também apresentaram problemas quanto à síntese, que não pôde ser avaliada dada a natureza da grade (avalia-se o uso do texto-insumo no critério Propósito. Essa proposta, de forma atípica, exigiu elementos do texto-insumo na construção do Gênero). Em relação à analogia, poucos candidatos estabeleceram contato com a fábula original, dificultando a construção de um parâmetro de correção. Além disso, inúmeros candidatos se deixaram conduzir pelos negritos do enunciado e comprometeram o fechamento de seu texto (ao tratarem da importância da leitura e não do conteúdo da fábula) ou o desenvolvimento do Gênero (ao “conversarem” com a comissão julgadora).

Após essas considerações, retomamos as perguntas que nortearam nossa pesquisa:

A prova de redação da Unicamp apresenta variações ou divergências em suas propostas (enunciados e textos) de uma edição para outra? De que natureza são essas variações?

A resposta é SIM.

As propostas variam não apenas de um ano para outro, mas na mesma edição do exame, visto que é solicitada a produção de dois textos distintos.

Mudam os gêneros solicitados, a natureza e o gênero dos textos que servem como insumo para a produção, os interlocutores, o grau de formalidade da linguagem exigida, o suporte, o conhecimento prévio e a autenticidade das tarefas, elementos observados em nosso protocolo de seleção das propostas analisadas. Esse tipo de variação, entretanto, é positivo, pois permite avaliar habilidades de leitura e de escrita distintas, que se complementam quando somados os dois textos produzidos.

No entanto, há variações negativas, como no caso da resenha solicitada em 2016, na qual observamos a presença de negritos que destacavam aspectos secundários, enunciados com excesso de informação e artificialidade da situação de produção. Todos esses aspectos comprometem a avaliação das habilidades do examinando, pois se deixa de avaliar as habilidades que eles têm de fato para verificar quais deles não se deixaram influenciar por problemas no enunciado.

Diferentes edições desse exame podem estar avaliando “habilidades diferentes” no que está sendo considerado “mesmo nível de proficiência”? Essas variações ameaçam o construto do exame?

A resposta é SIM. Podemos concluir que alunos com habilidades diferentes podem receber a mesma nota devido a inconsistências nos enunciados. No entanto, esses problemas observados na construção dos enunciados ou na seleção dos textos-insumo são minimizados pelo fato de haver sempre duas tarefas distintas. Concluímos que variações distintas não só ameaçam o construto do exame como também comprometem as representações dos examinandos sobre as tarefas, afetando seu desempenho, o processo de correção e os resultados da avaliação.

Uma das contribuições teóricas importantes do estudo foi mostrar a necessidade de maior explicitação do construto “integração da leitura e escrita” e que sua operacionalização nas tarefas na prova de redação do Vestibular Unicamp não introduza variações que possam comprometer esse construto, causando distorções nos resultados e, portanto, constituindo fontes de invalidade (MESSICK, 1989; SCARAMUCCI, 2011; SCARAMUCCI, 2017). O estudo, então, traz algumas evidências de validação no que diz respeito ao quesito “propostas” dos exames (enunciados/textos). Estudos de validação não têm

sido comuns no Brasil, ainda mais se considerarmos o conceito de validade contemporâneo (CHAPELLE, 1999; MESSICK, 1989; SCARAMUCCI, 2011), em que o foco é a validação do exame “em uso”. Os resultados teóricos e práticos serão de interesse não apenas nacional, mas também internacional, considerando o tópico abordado, o número de examinandos envolvidos e a relevância do exame para a sociedade.

As contribuições práticas decorrentes do estudo são relativas ao aprimoramento das propostas e, conseqüentemente, do exame e dos procedimentos de elaboração, correção e capacitação de corretores. De forma mais explícita, contribui também para o aprimoramento dos construtos avaliados e das grades de avaliação. Como se trata de um exame de alta relevância, em que decisões importantes são tomadas com base em seus resultados, as contribuições para a qualidade da avaliação e seus impactos na sociedade são grandes.

Podemos considerar uma limitação do estudo o fato de restringir-se a análises do produto dos examinandos (textos). Muito interessante e enriquecedor seria conduzir estudos futuros com foco nos processos/estratégias utilizados pelos examinandos/alunos para realizar uma tarefa integrada, podendo incluir, por exemplo: (1) seleção de ideias importantes a partir dos textos originais, (2) uso diferenciado de material de dois textos-insumo, (3) estilo de integração do material-fonte na escrita, (4) variáveis de cópia literal e paráfrases, (5) práticas de empréstimo do texto-insumo e seu efeito sobre a pontuação geral, (6) compreensão dos materiais-fonte, e (7) uso dos materiais-fonte para seleção de ideias, para apoiar opiniões, para apoio linguístico e para organizar o texto escrito.

Referências

ASENCIÓN-DELANEY, Y. Investigating the read-to-write construct. *Journal of English for Academic Purposes*, v. 7, n. 3, p. 140-150, jul. 2008.

CHAPELLE, C. A. Validity in language assessment. *Annual Review of Applied Linguistics*, v. 19, p. 254-272, 1999.

CHO, Y.; RIJMEN, F.; NOVÁK, J. Investigating the effects of prompt characteristics on the comparability of TOEFL-iBT integrated writing task. *Language Testing*, v. 30, n. 4, p. 513-534, 2013.

COMVEST. *Vestibular Unicamp*: Grade holística de correção da redação. Material de uso interno, 2010.

COMVEST. *Vestibular Unicamp*: redações 2013. Campinas: Editora da Unicamp, 2014.

_____. *Vestibular Unicamp*: redações 2016. Campinas: Editora da Unicamp, 2016.

CUMMING, A. Assessing integrated writing tasks for academic purposes: promises and perils. *Language Assessment Quarterly*, v. 10, n. 1, p. 1-8, 2013.

GEBRIL, A.; PLAKANS, L. Toward a transparent construct of reading-to-write tasks: the interface between discourse features and proficiency. *Language Assessment Quarterly*, v. 10, n. 1, p. 9-27, 2013.

KNOCH, U.; SITAJALABHORN, W. A closer look at integrated writing tasks: towards a more focused definition for assessment purposes. *Assessing Writing*, v. 18, n. 4, p. 300-308, out. 2013.

MESSICK, S. Meaning and values in test validation: the science and ethics of assessment. *Educational Researcher*, v. 18, n. 2, p. 5-11, mar. 1989.

PILEGGI, M. G. S. *Tarefas integradas nos exames de proficiência Celpe-Bras e TOEFL-iBT*. 2015. Dissertação (Mestrado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas,

Campinas. Disponível em: <http://repositorio.unicamp.br/bitstream/RE-POSIP/269574/1/Pileggi_MariaGabrielaSilva_M.pdf>. Acesso em: 14 out. 2017.

PLAKANS, L. Comparing composing processes in writing-only and reading-to-write test tasks. *Assessing Writing*, v. 10, p. 111-129, 2008.

SCARAMUCCI, M. V. R. A avaliação de habilidades integradas na parte escrita do exame Celpe-Bras. In: ALVAREZ, Maria Luisa Ortiz; GONÇALVES, Luis (Org.). *O mundo do português e o português no mundo afora: especificidades, implicações e ações*. Campinas: Pontes Editores, 2016. p. 391-425.

_____. Prova de redação nos vestibulares: educacionalmente benéfica para o ensino/aprendizagem da escrita? In: FLORES, Valdir do Nascimento (Org.). *A redação no contexto do Vestibular 2005: a avaliação em perspectiva*. Editora UFRGS, 2005, p. 37-57.

_____. Validade e consequências sociais das avaliações em contextos de ensino de línguas. *Lingvarvm Arena*, v. 2, p. 103-120, 2011.

_____. Relatório final de projeto de pesquisa intitulado “Diferentes propostas de produção escrita e seus efeitos no desempenho de alunos em três exames”. 2017 (não publicado).

SCARAMUCCI, M. V. R.; LOPES, R.; FRUNGILLO, M. *Proposta para a prova de redação do Vestibular Unicamp*. Documento interno, 2010.

SPIVEY, N. N.; KING, J. R. Readers and writers composing from sources. *Reading Research Quaterly*, v. 24, n. 1, p. 7-26, 1989.

WEIGLE, S. C.; YANG, W.; MONTEE, M. Exploring reading processes in an academic reading test using short-answer questions. *Language Assessment Quaterly*, v. 10, n. 1, p. 28-48, 2013;

YU, G. From integrative to integrated language assesement: are we there yet? *Language Assessment Quaterly*, v. 10, n. 1, p. 110-114, 2013.

Anexo I

Proposta adequada⁷

TEXTO 2

Imagine que, ao ler a matéria "Cães vão tomar uma 'gelada' com cerveja pet", você se sente incomodado por não haver nela nenhuma alusão aos possíveis efeitos que esse tipo de produto pode ter sobre o consumo de álcool, especialmente por adolescentes. Como **leitor assíduo**, você vem acompanhando o debate sobre o **álcool na adolescência** e decide escrever uma **carta** para a seção **Leitor** do jornal, criticando a matéria por não mencionar o problema do aumento do consumo de álcool.

Nessa carta, **dirigida aos redatores do jornal**, você deverá:

- fazer menção à matéria publicada, de modo que mesmo quem não a tenha lido entenda a importância da crítica que você faz;
- fundamentar a sua crítica com dados apresentados na matéria "Vergonha Nacional", reproduzidos adiante.

Atenção: ao assinar a carta, use apenas as iniciais do remetente.

Cães vão tomar uma "gelada" com cerveja pet

Produto feito especialmente para cachorros chega ao mercado nacional em agosto

Nada é melhor que uma cervejinha depois de um dia de cão.

Agora eles, os cães, também vão poder fazer jus a essa máxima. No mês de agosto chega ao mercado a Dog Beer, cerveja criada especialmente para os amigos de quatro patas. "Quem tem bicho de estimação gosta de dividir o prazer até na hora de comer e beber", aposta o empresário M. M., 47, dono da marca.

Para comemorar a final da Libertadores, a executiva A. P. C., 40, corintiana roxa, quis inserir Manolito, seu labrador, na festa.

"Ele tomou tudo. A cerveja é docinha, com fundinho de carne", descreve.

Uniformizado, Manolito não só bebeu a gelada durante o jogo contra o Boca Juniors como latiu sem parar até o fim da partida.

Desenvolvida pelo centro de tecnologia e formação de cervejeiros do Senai, no Rio de Janeiro, a bebida canina é feita à base de malto e extrato de carne; não tem álcool, lúpulo, nem gás carbônico.

O dono da empresa promete uma linha completa de "petscos líquidos", que inclui suco, vinho e champanhe.

A lista de produtos humanos em versões animais não para de crescer.

Já existem molhos, tempero para ração e até patê.

O sorvete Ice Pet é uma boa opção para o verão. A sobremesa tem menos lactose, não tem gorduras nem açúcar.

(Adaptado de Ricardo Bunduky, Folha de São Paulo, São Paulo, 22 julh.2012. Cotidiano 3 p.)

Vergonha Nacional

As décadas de descumprimento da lei (...) contribuíram para que os adultos se habituassem a ver o consumo de bebidas entre adolescentes como "mal menor", comparado aos perigos do mundo. (...) Um estudo publicado pela revista *Drugs and Alcohol Dependence* ouviu 15.000 jovens nas 27 capitais brasileiras. O cenário que emerge do estudo é alarmante. Ao longo de um ano, um em cada três jovens brasileiros de 14 a 17 anos se emborrou ao menos uma vez. Em 54% dos casos mais recentes, isso ocorreu na sua casa ou na de amigos ou parentes. Os números confirmam também a leniência com que os adultos encaram a transgressão. Em 17% dos episódios, os menores estavam acompanhados dos próprios pais ou de tios.

Resultados da pesquisa realizada com 15.000 jovens de 14 a 17 anos nas 27 capitais brasileiras

Quantas vezes se emborrou		Onde ficou emborrou (na última vez em que bebeu)		Com quem bebeu (na última vez em que bebeu)	
Nenhuma vez	12%	Bar	35%	Amigos	50%
Uma vez na vida	35%	Casa de amigos	30%	Irmãos e primos	26%
Ao menos uma vez no último ano	32%	Casa de parentes	13%	Pais ou tios	17%
Ao menos uma vez no último mês	21%	Própria casa	11%	Namorado	5%
		Festas ou praia	11%	Sozinho	2%

(Adaptado de Revista Veja, Editora Abril, São Paulo, nº 26, 11 julh. 2012, p. 61-62.)

7 Disponível em: <http://www.comvest.unicamp.br/vest2013/F1/f12013QZ.pdf>

Anexo II

Proposta potencialmente inadequada⁸

UNICAMP
vestibular
2016

REDAÇÃO

TEXTO 1

Você é um estudante universitário que participará de um **concurso de resenhas**, promovido pelo Centro de Apoio ao Estudante (CAE), órgão que desenvolve atividades culturais em sua Faculdade. Esse concurso tem o objetivo de **estimular a leitura** de obras literárias e **ampliar o horizonte cultural** dos estudantes. A **resenha** será lida por uma **comissão julgadora** que deverá selecionar os dez melhores textos, a serem publicados. Você escolheu resenhar a fábula de La Fontaine transcrita abaixo. Em seu texto, você deverá incluir:

- uma síntese da fábula, indicando os seus elementos constitutivos;
- a construção de uma situação social análoga aos fatos narrados, que envolva um problema coletivo;
- um fechamento, estabelecendo relações com a temática do texto original.

Seu texto deverá ser escrito em **linguagem formal**, deverá indicar o **título da obra** e ser assinado com um **pseudônimo**.

A Deliberação Tomada pelos Ratos

Rodilardo, gato voraz,
aprontou entre os ratos tal matança,
que deu cabo de sua paz,
de tantos que matava e guardava na pança.
Os poucos que sobraram não se aventuravam
a sair dos buracos: mal se alimentavam.
Para eles, Rodilardo era mais que um gato:
era o próprio Satã, de fato.
Um dia em que, pelos telhados,
foi o galante namorar,
aproveitando a trégua, os ratos, assustados,
resolveram confabular
e discutir um modo de solucionar
esse grave problema. O decano, prudente,
definiu a questão: simples falta de aviso,
já que o gato chegava, solerte. Era urgente
amarrar-lhe ao pescoço um guizo,
concluiu o decano, rato de juízo.
Acharam a ideia excelente,
e aplaudiram seu autor. Restava, todavia,
um pequeno detalhe a ser solucionado:
quem prenderia o guizo – e qual se atreveria?
Um se esquivou, dizendo estar muito ocupado;
Outro alegou que andava um tanto destreinado
em dar laços e nós. E a bela ideia
teve triste final. Muita assembleia, ao fim nada decide – mesmo sendo de frades
ou de veneráveis abades...

Deliberar, deliberar ...
conselheiros, existem vários;
mas quando é para executar,
onde estarão os voluntários?

(*Fábulas de La Fontaine*. Tradução de Milton Amado e Eugênia Amado. Belo Horizonte: Itatiaia, 2003, p. 134-136.)

Glossário

Abade: superior de ordem religiosa que dirige uma abadia.
Frade: indivíduo pertencente a ordem religiosa cujos membros seguem uma regra de vida e vivem separados do mundo secular.
Decano: o membro mais velho ou mais antigo de uma classe, assembleia, corporação, etc.
Guizo: pequena esfera de metal com bolinhas em seu interior que, quando sacudida, produz um som tilintante.
Solerte: engenhoso, esperto, sagaz, artiloso, arguto, astucioso.

⁸ Disponível em: http://www.COMVEST.unicamp.br/vest_anteriores/vest2013.html

Matilde Virginia Ricardi Scaramucci

Pós-doutora em Linguística pela Universidade de Melbourne, Austrália.

Professora da Unicamp, Campinas (SP), Brasil;

matilde@scaramucci.com.br

Luciana Amgarten Quitzau

Mestre em Linguística Aplicada pela Unicamp e doutoranda do mesmo programa.

Pesquisadora na Coordenação Acadêmica, Unicamp, Campinas (SP), Brasil;

lu_amgarten@yahoo.com.br

Artigo submetido em 23/10/2017

Aprovado em 30/05/2018

AUTOAVALIAÇÃO INSTITUCIONAL NA POLÍTICA DO SINAES: CONCEPÇÃO, DESAFIOS E POSSIBILIDADES DE AVANÇOS

INSTITUTIONAL SELF-EVALUATION IN THE SINAES POLICY: CONCEPTION,
CHALLENGES AND POSSIBILITIES OF ADVANCES

AUTOAVALACIÓN INSTITUCIONAL EN LA POLÍTICA DEL SINAES:
CONCEPCIÓN, DESAFÍOS Y POSIBILIDADES DE AVANCES

José Vieira de Sousa

RESUMO

A reflexão proposta neste artigo problematiza a autoavaliação institucional no âmbito da política do Sistema Nacional de Avaliação da Educação Superior (Sinaes), situando-a tanto em relação à perspectiva legal que a institui como motriz desse sistema quanto na dinâmica que ela tem assumido na avaliação desse nível educacional, no Brasil. Para isso, focaliza os desafios e avanços subjacentes à atuação das Comissões Próprias de Avaliação, traduzidos nos relatórios de autoavaliação. Discute-se a natureza e a finalidade da autoavaliação, ressaltando a polissemia que envolve o conceito de qualidade na educação superior. A análise empreendida conclui que, devido à importância que possui, a autoavaliação pode se configurar como instrumento de gestão pedagógica, considerando o uso dos seus resultados, em relação ao trabalho que as instituições de educação superior realizam. Aponta também alguns desafios relativos à implementação da política do Sinaes, ressaltando que, à proporção que as referidas instituições não consolidam sua autoavaliação, elas poderão revelar dificuldades práticas, em relação à definição estratégica de suas ações futuras. A não centralidade da autoavaliação institucional tem na ausência do próprio Estado, no que tange à ausência de acompanhamento sistemático e cíclico do trabalho das comissões antes citadas, uma das principais razões para o insuficiente reconhecimento desse componente avaliativo no âmbito do Sinaes. O que

se tem verificado, na prática, é a postagem dos relatórios de autoavaliação pelas instituições para atender ao ciclo avaliativo, porém sem retorno efetivo do Estado em relação a esse componente da política de avaliação da educação superior brasileira, em vigência desde 2004.

Palavras-chave: educação superior; Sinaes; autoavaliação.

ABSTRACT

The reflection proposed in this article problematizes the institutional self-assessment within the scope of the National System of Evaluation of Higher Education (Sinaes) policy, situating it in relation to both the legal perspective that establishes it as the motive of this system and the dynamics that it has assumed in the evaluation of this educational level, in Brazil. In order to do so, it focuses on the challenges and advances underlying the performance of the Evaluation Committees, translated into the reports of institutional self-evaluation. It discusses the nature and purpose of self-assessment, highlighting the polysemy that involves the concept of quality in higher education. The analysis carried out concludes that, due to its importance, self-assessment can be configured as an instrument of pedagogical management, considering the use of its results, in relation to the work that higher education institutions perform. It also points out some challenges related to the implementation of the Sinaes policy, emphasizing that to the extent that these institutions do not consolidate their self-assessment, they may reveal practical difficulties in relation to the strategic definition of their future actions. The absence of systematic and cyclical monitoring of the work of the aforementioned commissions is one of the main reasons for the insufficient recognition of this evaluation component within Sinaes. What has been verified in practice is the posting of self-evaluation reports by the institutions to attend the evaluation cycle, but without effective return from the State in relation to this component of the evaluation policy of Brazilian higher education, in force since 2004.

Keywords: higher education; Sinaes; self-evaluation.

RESUMEN

El artículo tiene por objetivo reflexionar sobre la autoevaluación institucional como componente del Sistema Nacional de Evaluación de la Educación Superior (Sinaes), problematizándola en una doble dirección. Por un lado, destaca la centralidad asegurada a ese componente por la Ley nº 10.861/2004, a medida que lo concibe como motivo del proceso de evaluación y, por otro, algunos impasses presentes en su materialización frente a la política del Sinaes, así como posibilidades de avances del trabajo de las Comisiones Propias de Evaluación, traducidas en los informes de autoevaluación institucional. Discute la naturaleza y finalidad de la autoevaluación, resaltando la polisemia que involucra el concepto de calidad en la educación superior. El análisis emprendido concluye que, debido a la importancia que posee, la autoevaluación puede configurarse como instrumento de gestión pedagógica, considerando el uso de sus resultados, en relación al trabajo que las instituciones de educación superior realizan. En el presente trabajo se analizan los resultados obtenidos en el análisis de los resultados obtenidos en el estudio. La no centralidad de la autoevaluación institucional tiene en la ausencia del propio Estado, en lo que se refiere a la ausencia de seguimiento sistemático y cíclico del trabajo de las comisiones antes citadas, una de las principales razones para el insuficiente reconocimiento de ese componente evaluativo en el ámbito del Sinaes. Lo que se ha comprobado, en la práctica, es la postura de los informes de autoevaluación por las instituciones para atender el ciclo de evaluación, pero sin retorno efectivo del Estado en relación a ese componente de la política de evaluación de la educación superior brasileña, en vigencia desde 2004.

Palabras clave: educación superior; Sinaes; autoevaluación.

Introdução

O objetivo deste artigo é refletir sobre a autoavaliação institucional como componente do Sistema Nacional de Avaliação da Educação Superior (Sinaes), problematizando-a em uma dupla direção. De um lado, o texto enfatiza a centralidade assegurada a esse componente pela Lei nº 10.861, de 14 de abril de 2004 (BRASIL, 2004a), na medida em que o concebe como motriz do processo avaliativo e, por outro lado, destaca alguns desafios

e impasses presentes em sua materialização face à política do Sinaes, bem como possibilidades de avanços do trabalho das Comissões Próprias de Avaliação (CPAs), cuja tradução ocorre nos relatórios de autoavaliação institucional.

Por meio das Portarias MEC/SESu nº 11, de 28 de abril de 2003 (BRASIL, 2003a), e nº 19, de 27 de maio de 2003 (BRASIL, 2003b), deu-se início à geração do Sinaes com a realização de estudos pela Comissão Especial de Avaliação da Educação Superior (CEA). A finalidade precípua dessa Comissão foi propor mudanças relevantes nos procedimentos de avaliação da educação superior brasileira. De acordo com o art. 1º da Portaria nº 11/2003, o grupo de trabalho que formou a comissão teve como objetivos “analisar, oferecer subsídios, fazer recomendações, propor critérios e estratégias para a reformulação dos processos e políticas de avaliação da Educação Superior e elaborar a revisão crítica dos seus instrumentos, metodologias e critérios utilizados”.

A partir da ação da CEA foi instituído o Sinaes pela Lei nº 10.861/2004 (BRASIL, 2004a), que define como seu objetivo assegurar o processo nacional de avaliação das instituições de educação superior (IES), dos cursos de graduação e do desempenho acadêmico de seus estudantes (art. 1º). Em decorrência do seu objetivo mais amplo, o Sinaes tem como finalidades: (i) a melhoria da qualidade da educação superior; (ii) a orientação da expansão da sua oferta; (iii) o aumento permanente da sua eficácia institucional e da sua efetividade acadêmica e social; e (iv) a intensificação do aprofundamento dos compromissos e das responsabilidades sociais das IES por meio da valorização de sua missão pública, bem como dos valores democráticos, do respeito à diferença e à diversidade, da afirmação da autonomia e da identidade institucional.

Essa mesma lei determinou a criação da Comissão Nacional de Avaliação da Educação Superior (Conaes), a qual tem como principal função a coordenação e a supervisão do Sinaes. De acordo com o art. 6º, inciso I, da lei, uma das atribuições dessa Comissão é “propor e avaliar as dinâmicas, procedimentos e mecanismos da avaliação institucional, de cursos e de desempenho dos estudantes”. Por sua vez, a Portaria nº 2.051, de 9 de julho de 2004 (BRASIL, 2004b), deu início à regulamentação dos procedimentos de avaliação do Sinaes.

Estruturando-se a partir da concepção (dimensão teórica) e do sentido (dimensão prática) os quais a autoavaliação vem assumindo no contexto do Sinaes, o artigo está organizado em três partes: introdução, desenvolvimento e conclusão. A primeira parte explicita o objetivo do texto e as bases legais que dão origem ao Sinaes, como política de avaliação da educação superior brasileira. A segunda está subdividida em três seções: a inicial procede à contextualização dos três componentes do Sinaes, destacando a natureza e a finalidade da autoavaliação institucional; a seção seguinte problematiza o conceito de qualidade na educação superior, ressaltando a polissemia que o envolve; a terceira seção trata da materialidade da qualidade traduzida nos relatórios de autoavaliação produzidos pelas CPAs; e a última parte do artigo corresponde à conclusão, o que permite retomar algumas ideias enfatizadas nas partes que a antecedem.

O Sinaes como política pública

Na atualidade, o Sinaes figura como a principal ferramenta de avaliação da educação superior brasileira, passando por medidas que visam ao seu aprimoramento constante. Na condição de um processo global de avaliação, sua efetivação requer a integração das instâncias e das pessoas responsáveis pelas ações e pelos processos que o configuram como tal. Sua estruturação tomou como referência uma tríade de componentes: avaliação das instituições (interna e externa), dos cursos e do desempenho dos estudantes, com objetivo de promover a articulação e a integração desses componentes.

Componentes do Sinaes: finalidades e alcance

No que tange à *avaliação das instituições*, o Sinaes desenvolve dois processos complementares: interno e externo. Este último é realizado por comissões externas, compostas por especialistas, por meio de avaliações que têm como objetivo verificar os dados informados pela IES previamente no sistema e-MEC.¹ As comissões observam aspectos diversos relacionados a dez dimensões estabelecidas na Lei nº 10.861/2004. No caso da *avaliação*

¹ Em termos operacionais, o sistema e-MEC constitui uma ferramenta que possibilita ao Inep implementar critérios de designação, obedecendo aos perfis dos avaliadores requeridos a cada processo de avaliação e de forma automática, por meio de uma escolha aleatória desses perfis para que os avaliadores componham comissões para fazer visitas *in loco*.

interna, essa lei institui a criação da CPA em cada IES pública ou privada. Embora atue de forma autônoma no processo de autoavaliação institucional, essa Comissão estrutura seu trabalho à luz de um roteiro sugerido pela Conaes e pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Isso estabelece os critérios e as dimensões a serem considerados.

Quanto ao segundo componente do Sinaes – *avaliação dos cursos de graduação* – de acordo com o art. 4º da Lei nº 10.861/2004, seu objetivo é identificar as condições de ensino oferecidas aos estudantes pela instituição. Tal avaliação acontece tomando como referência três grandes dimensões: o perfil do corpo docente, a infraestrutura e organização didático-pedagógica e os seus respectivos indicadores. Ao contemplar vários elementos do curso, essa avaliação acaba por promover o seu diagnóstico.

O terceiro componente – *Exame Nacional de Desempenho dos Estudantes (Enade)* – propõe aferir o desempenho dos discentes em relação aos conteúdos programáticos estabelecidos nas diretrizes curriculares de cada curso de graduação. Para tanto, focalizam-se as habilidades e as competências dos estudantes frente às exigências da evolução do conhecimento e da realidade brasileira e mundial no que diz respeito aos temas ligados às suas profissões, conforme o art. 5º, § 1º, da Lei nº 10.861/2004. A aplicação do Enade é feita periodicamente, tendo sido admitido, até 2008, o percentual amostral aos estudantes de todos os cursos de graduação, ao final do primeiro e do último ano de curso. A partir de 2009, sua realização passou a ser censitária, alcançando a todos os estudantes que possuem condições de participar dele, nos termos definidos pela lei. O art. 23 da mesma lei define como objetivo do Enade “acompanhar o processo de aprendizagem e o desempenho dos estudantes em relação aos conteúdos programáticos previstos nas diretrizes curriculares do respectivo curso de graduação” (BRASIL, 2004a).

Contextualizados os três componentes do Sinaes, enfatizaremos a autoavaliação institucional, tendo em vista que ela constitui o foco da discussão proposta no artigo. Preliminarmente, destaca-se que cabe à CPA conduzir os processos de avaliação interna da IES, a sistematização e a prestação de informações diversas ao Inep. Esses elementos concorrem para o reconhecimento do principal objetivo da proposta dessa avaliação: possibilitar a cada instituição

o autoconhecimento, a análise de seus processos e, conseqüentemente, a reflexão e contextualização do seu trabalho para as diversas avaliações externas.

Considerando a importância e a validade para o aperfeiçoamento do trabalho da instituição em suas várias dimensões, a autoavaliação amplia a visão da comunidade interna sobre suas potencialidades e suas fragilidades para a tomada de decisão (BELLONI, 1999; DIAS SOBRINHO, 2008). Desse ponto de vista, essa avaliação pode levar a IES a reconhecer seus pontos fortes e fracos, a planejar ações e a definir mecanismos de correção e de transformação, o que visa melhorar a qualidade de suas atividades. Essa visão formativa é expressa no documento que propõe a instituição do Sinaes (INEP, 2009, p. 105):

Entre os objetivos da avaliação contam-se o de conhecer as fortalezas e os problemas da instituição [...]. Não basta levantar as deficiências. É também muito importante identificar as qualidades e aspectos fortes da IES. Quanto aos problemas e carências, além da verificação e da constatação, é importante identificar as causalidades, explicitar as possibilidades reais para a superação e estabelecer as ações adequadas e os meios para a transformação desejada.

Na prática, é por meio da autoavaliação que um curso ou uma instituição se analisam internamente a fim de melhor compreender como realizam suas práticas e como se organizam, além de procurar sistematizar e analisar dados e informações que ajudam a identificar práticas exitosas e outras que merecem ser repensadas para evitar futuros equívocos. Isso oportuniza à IES a busca por seu autoconhecimento, o qual, para Belloni, significa “identificar os acertos e as ineficiências, as vantagens/potencialidades e as dificuldades; envolver um processo de reflexão sobre as razões, as causas das situações positivas e das insuficiências” (1996, p. 8). Na visão formativa, ao constatar suas fragilidades e corrigi-las pela autoavaliação, a IES estará passando por um processo de autorregulação, de maneira a alcançar, com mais qualidade, seus objetivos e suas missões (INEP, 2009).

Ampliando o escopo da discussão sobre o tema avaliação, como toda política pública, o Sinaes é dotado de grande complexidade, situando-se em relação a um campo igualmente complexo, heterogêneo e revelador de grandes contradições – o da educação superior brasileira. Tomando

como referência esse cenário, pesquisadores, como, por exemplo, Barreyro (2008), Dias Sobrinho (2010), Leite (2008), Peixoto (2011) e Sousa (2016), têm formulado críticas à política avaliativa instituída por esse sistema.

Todavia, ainda que se considere a pertinência do teor das críticas levantadas pelos autores mencionados e por outros, há que se levar em conta, também, a posição assumida por outra parcela de estudiosos para os quais a mesma política vem se traduzindo em avanços relevantes. Subjaz à defesa desse ponto de vista a tese de que o Sinaes é orientado por uma perspectiva paradigmática emancipatória/formativa (POLIDORI; ARAÚJO; BARREYRO, 2006). Na visão de autores que fazem parte do segundo grupo, a oposição da política avaliativa em questão é feita diretamente à visão que promove e/ou estimula *rankings* e competitividade entre as IES. Em uma direção análoga, Griboski e Fernandes (2012) sustentam que o Sinaes se mostrou como um:

[...] meio eficaz de garantir acesso às informações que sinalizam qualidade [e que] se formativa, diagnóstica, emancipatória, imposta compulsoriamente pelo Estado ou não, a avaliação possibilitou traçar o panorama da Educação Superior, por meio de uma fotografia da situação educacional, mais aproximada da realidade (p. 108).

Com efeito, à proporção que se insere no campo da educação superior brasileira – objeto de disputas de grupos que possuem interesses distintos em relação a ele –, o Sinaes permite uma diversidade de olhares sobre sua política avaliativa. Partindo dessa premissa, é importante aprofundar o debate sobre a concepção e a natureza da autoavaliação, bem como sinalizar desafios e contradições revelados por esse componente no contexto do sistema avaliativo em questão.

Compreendida no escopo de uma política pública, a avaliação da educação superior deve levar em conta os impactos, os objetivos e as diretrizes para determinar a qualidade desse nível educacional. Ela deve servir, igualmente, de permanente aprendizado para todos os envolvidos que detêm o conhecimento sobre a sua relevância, a eficácia e a sustentabilidade das ações as quais esse tipo de avaliação propõe. Para tanto, é preciso garantir a articulação entre seus marcos teóricos, políticos e técnicos. Segundo Dias Sobrinho (2010), esses marcos contribuem para compreender

que a avaliação possui problemas, limites, incompletudes, variedade de concepções e um expressivo número de dificuldades insolúveis. Todavia, ainda conforme o autor, “o propósito de obter rigor técnico e facilidades operacionais não deve justificar as soluções simplistas, que reduz todo o sistema de avaliação a uma ferramenta de medição do rendimento acadêmico para efeito de regulação” (ibid., p. 223).

É nesse contexto interpretativo que defendemos a possibilidade da autoavaliação vir a ser um importante instrumento de gestão da IES. Entretanto, para isso, não é suficiente que cada instituição se limite a atender aos dispositivos legais e burocráticos, mas que busque, sobretudo, a contínua melhoria da qualidade dos serviços que presta à sociedade. Assim, a autoavaliação amplia suas chances de se constituir, de fato, em uma ferramenta dinâmica, interativa e processual para o trabalho da IES. Trata-se, portanto, de um processo em contínua construção, perspectiva que mantém relação direta com a tese política de que “não se pode conceber o mundo como um conjunto de coisas acabadas, mas como um conjunto de processos” (MARX; ENGELS, 1963, p. 195).

A preocupação com a autoavaliação no contexto do Sinaes ganha densidade na medida em que é concebida como uma proposta avaliativa que, orientada por princípios formativos, pressupõe a integração de suas partes e dos instrumentos que visam sua operacionalização. Por outro lado, é preciso alertar para o fato de que a própria incompreensão, acerca do que vem a ser a autoavaliação em suas finalidades formativas, pode dificultar uma concepção ampliada a seu respeito, pautada, sobretudo, nos pressupostos teórico-metodológicos assumidos pelo Sinaes.

Para finalizar esta seção do texto, convém lembrar que o cenário atual demanda reflexões e decisões sobre como o Sinaes precisa avançar visando a efetiva integração de seus componentes e de suas várias dimensões. Defendemos essa ideia na perspectiva de que a avaliação se configure, de fato, em um empreendimento sistemático que tenha por finalidade precípua a compreensão global da IES. Em decorrência, a autoavaliação pode vir a ser reconhecida, nesse processo, como um componente que permite cada IES se autoconhecer, considerando a totalidade de suas práticas acadêmicas.

Qualidade: ponto de partida e de chegada ou pomo da discórdia na educação superior?

Tem sido recorrente, na literatura especializada, a ausência de consenso em relação ao que se compreende por qualidade, notadamente quando o debate do tema recai no complexo campo da educação superior. Todavia, talvez este fato seja revestido de positividade, na medida em que a opção do próprio governo, em não definir qualidade de forma cabal, permita comportar, dentro do termo, um variado conjunto de práticas. Concretamente, essas práticas se expressam em diferentes graus, que se ajustam, entre outros, ao perfil institucional, à vocação, à geografia e às questões sociais que envolvem as IES. Não é por acaso que a escala adjetiva/adverbial, adotada nos formulários de avaliação da graduação (Inep), considera a qualidade moderada em níveis de 1 a 5, ao passo que, na pós-graduação, a métrica da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) a focaliza em níveis distintos – de 1 a 7.

Por outro lado, há também convergência na mesma literatura para a ideia de que o conceito de qualidade é dotado de grande polissemia e, ao mesmo tempo, é objeto de disputa por diferentes grupos que desejam ver a ele associados seus interesses e suas aspirações, quando associada ao tema investigado neste artigo – autoavaliação institucional na política do Sinaes – tal polissemia é intensificada, pois “o entendimento de qualidade é inexoravelmente subjetivo, porque depende fundamentalmente das concepções de mundo e de Educação Superior de quem o emite” (BERTOLIN, 2007, p. 156). Mantendo sintonia com essa ideia, Sousa (2009, p. 244) registra que:

[...] as definições para a palavra qualidade são inúmeras e mutáveis. Por isso, a transformação é fator essencial para que se compreenda, de forma mais contextualizada e crítica, o seu percurso histórico-social e como ele se deixa construir. Esta preocupação ganha relevo, particularmente, pela carga subjetiva e polissêmica que a palavra qualidade envolve, quando aplicada à educação.

Mesmo sendo assim, não só as IES devem estar preparadas para compreender e administrar, internamente, a polissemia do conceito em questão, alocando seus planos de desenvolvimento, as metas e os objetivos ao sentido de qualidade que melhor se ajusta ao seu perfil e missão. Mas

também deve o governo, por meio de seus colaboradores, seus agentes internos e seus avaliadores, olhar as IES com as lentes ajustadas, rejeitando um modelo padronizado de qualidade. Entretanto, ao que tudo indica, esse tema ainda não tem sido contemplado na agenda das instâncias governamentais envolvidas com a avaliação da educação superior, no país, pelo menos na intensidade e na abrangência que merece.

De forma constante, a melhoria da qualidade da educação superior tem alimentado intensamente os debates promovidos por diferentes segmentos sociais. Para ilustrar isso, basta ver a quantidade de projetos de lei com vistas a alterar a Lei de Diretrizes e Bases (LDB) nº 9.394/1996 (BRASIL, 1996) e a Lei nº 10.861/2004, as quais instituiu o Sinaes, ou, até mesmo, observar o grande número de emendas ao Projeto de Lei nº 4.372/2012 (BRASIL, 2012), que propõe a criação do Instituto Nacional de Supervisão e Avaliação da Educação Superior (Insaes) e dá outras providências. Fatos semelhantes a esses comprovam como, cada vez mais, a qualidade ganha centralidade e importância no âmbito da avaliação da educação superior como política pública. Além disso, corrobora-se a tese sustentada por Azevedo (2008) de que toda política pública tem sua razão de ser quando focaliza um problema que tem relevância para a sociedade. Na perspectiva da autora, toda política pública – como, por exemplo, avaliação da educação superior – é formulada tendo como ponto de partida um problema que passa a ser amplamente debatido pela sociedade que, por sua vez, exige a atuação do Estado. Para ela, “pode-se afirmar que um setor ou uma política pública para um setor, constitui-se a partir de uma questão que se torna socialmente problematizada” (ibid., p. 61).

Retomando a questão conceitual da qualidade, Acioli e Oliveira (2014, p. 243), defendem que os embates teóricos e práticos que revestem esse termo na educação superior permitem constatar que este:

[...] pode ser traduzido a partir de uma visão considerada tradicional, em que aspectos como infraestrutura, planos de ensino, corpo docente entre outros, são considerados como elementos decisórios para qualificar uma IES. Já em outra perspectiva, qualidade é entendida em função da capacidade de ajustamento das IES às imposições dos agentes externos. Enquanto, em uma visão mais social, a qualidade passa pelo atendimento equitativo a todos que a requerem.

Em uma direção que preserve sintonia com as diretrizes da avaliação proposta pelo Sinaes, é imprescindível a defesa da qualidade na perspectiva de algo que não se esgote nos resultados e, assim, possa contribuir para a construção de uma proposta emancipatória. Isso implica a produção de ideias e de concepções avaliativas, bem como a definição de objetivos, de metas, de metodologias em torno de um conjunto que, além de se mostrar coeso, coerente e consistente, precisa articular teoria e prática. Essa indução à criatividade e à proatividade, no encaminhamento de propostas de autoavaliação, é, a nosso ver, para onde deve caminhar o modelo de avaliação, com forte incentivo do Inep e da Conaes, fomentando as ações de autoavaliação e de consciência de que esse empoderamento minimiza a necessidade de supervisão e traz mais autonomia às IES. Oportunamente, lembramos que isso não é disposição ou garantia decorrente do Sinaes, visto que o art. 54, § 2º, da LDB nº 9.394/1996, já previa que “as atribuições de autonomia universitária poderão ser estendidas as instituições *que comprovem alta qualificação para o ensino ou para a pesquisa, com base em avaliação realizada pelo Poder Público*” (BRASIL, 1996, grifo nosso).

Os relatórios das CPAs e a dimensão da materialidade da qualidade

Considerando a discussão feita na seção anterior, é preciso reconhecer, também, a dimensão da materialidade da qualidade, entendida como a necessidade de que os resultados, obtidos na autoavaliação institucional, possam provocar transformações na gestão da IES em uma dupla direção: (i) de um lado, provocando mudanças, a partir da readequação ou revisão das estruturas já existentes; e (ii) por outro, implicando a necessidade de mais investimentos para superar deficiências detectadas pelo processo avaliativo realizado. Nesse contexto, é imprescindível o engajamento dos gestores no processo, bem como a compreensão da necessidade da tomada de decisão para atender às duas situações antes apontadas. Trata-se, portanto, de os dirigentes ouvirem o que têm a dizer os resultados da autoavaliação nas várias dimensões do Sinaes. Entretanto, nesse ponto, ganha relevo, especialmente, uma contradição: tudo indica que, na realidade atual, os relatórios de autoavaliação parecem ser endereçados apenas ao governo. Por isso, é preciso não perder de vista que a perspectiva de futuro da autoavaliação requer efetivamente contemplar, também, como destinatário, os próprios dirigentes das IES.

Do ponto de vista da política do Sinaes, a ideia norteadora da autoavaliação institucional é que os relatórios, produzidos pelas CPAs, revelem esforços, visando fazer uma análise interna das práticas da IES na perspectiva de vislumbrar o que deseja ser, bem como as atividades que efetivamente desenvolve. Além disso, esses documentos deveriam apontar como se dá sua organização, sua administração, sua representatividade e, fundamentalmente, como cada IES age frente aos desafios implícitos à gestão da educação superior, sendo isso expresso, respeitando a identidade e a vocação de cada instituição, em níveis diversificados. Essa variação decorre de fatores diversos, como, por exemplo, organização acadêmica, categoria administrativa e, notadamente, a própria vivência de cada IES, em termos do seu percurso na autoavaliação.

À luz do que defende o Sinaes para a avaliação da educação superior brasileira, há a expectativa de que os relatórios de autoavaliação possam contribuir para que cada IES busque “se autoconhecer com vistas ao aperfeiçoamento de suas atividades acadêmicas, objetivando melhorar a qualidade educativa e alcançar relevância social” (AUGUSTO; BALZAN, 2007, p. 608). Para tanto, supõe-se que o trabalho da CPA seja convergente para a ideia de expressar coerência com o sentido da autoavaliação proposta que orienta a autoavaliação no escopo mais amplo da política do Sinaes.

Na reflexão realizada neste artigo, é importante considerar, também, que o processo, rumo à elaboração dos relatórios parciais e integrais da autoavaliação, pressupõe a participação dialogada entre a CPA e os demais segmentos das instituições, momento em que a Comissão colhe dados sobre metas e ações, potencialidades, fragilidades etc. Todavia, na produção desses documentos, convém que eles não tenham a mão de nenhum gestor da IES ou de seus mantenedores. Afinal, uma vez finalizados os relatórios, o canal de comunicação é entre CPA e governo, ou entre aquela e as Comissões Externas designadas pelo Inep para proceder às avaliações *in loco*.

Não é por acaso que, na Lei nº 10.861/2004, o Congresso Nacional garantiu que as CPAs fossem autônomas em relação às demais instâncias das IES, com a atribuição de condução dos processos de avaliação internos da instituição, de sistematização e de prestação das informações solicitadas pelo Inep. É importante que se perceba também que, embora as referidas Comissões colham e analisem informações dentro da IES, dialogando

com os demais setores, seu relatório final é dirigido ao governo, como já registrava o art. 13, §1º, da Portaria do MEC nº 2.051/2004 (BRASIL, 2004b).

Por sua vez, o art. 17-K da Portaria Normativa nº 40/2007 (BRASIL, 2007), republicada em 2010, diz que tais relatórios devem estar disponíveis aos agentes do governo, no caso as Comissões Externas de Avaliação:

Art. 17-K – Deverão estar disponíveis para análise pela Comissão de Avaliação previamente à realização da visita, além do formulário eletrônico de avaliação, outros documentos, que permitam considerar a instituição ou curso no conjunto, tais como: I – relatórios parciais e finais do processo de autoavaliação da instituição.

Enquanto isso, o art. 61 define como os relatórios serão encaminhados ao Ministério da Educação:

Art. 61-D – Será mantido no cadastro e-MEC, junto ao registro da instituição, campo para inserção de relatório de autoavaliação, validado pela CPA, a ser apresentado até o final de março de cada ano, em versão parcial ou integral, conforme se trate de ano intermediário ou final do ciclo avaliativo.

Cabe lembrar que a Portaria nº 92/2014² (BRASIL, 2014) prevê que as IES devam produzir, a partir de sua publicação, um Relato Institucional no Eixo Planejamento e Avaliação, documento considerado uma inovação no instrumento de Avaliação Institucional Externa. Nele deverá conter um relato do Plano de Desenvolvimento Institucional (PDI) e, especialmente, uma síntese histórica dos processos de avaliação interna e externa da IES e do seu planejamento. Deve explicitar, igualmente, as ações acadêmico-administrativas decorrentes dos resultados das avaliações, conforme o inciso VIII do art. 3º da Lei nº 10.861/2004. Desse ponto de vista, o referido relato pode contribuir significativamente em uma dupla direção: na discussão da própria concepção de qualidade expressa nas ações da IES e na

2 Ao direcionar o atendimento às exigências contidas na realização da avaliação institucional, a Portaria nº 92, de 31 de janeiro de 2014, revogou a Portaria nº 1.016, de 30 de outubro de 2007, que sancionou o Instrumento de Avaliação elaborado pelo Inep para credenciamento de novas IES de Sinaes; e a Portaria nº 1.264, de 17 de outubro de 2008, que aprovou o Instrumento de Avaliação Externa de IES desse sistema.

consideração de que a tomada de decisão interna tem a ver com assumir o que seja a qualidade mais ajustada ao perfil institucional.

Diante do exposto até aqui, mais uma vez, torna-se necessário ressaltar a importância de os relatórios das CPAs explicitarem o conceito de qualidade que norteia as práticas da IES, de maneira que este não se relacione somente a determinadas estratégias de controle e regulação exercidas pelo Estado. Na análise crítica que fazem do tema, Tavares, Peixoto e Correia (2014) ressaltam que vem ganhando terreno, no campo da educação superior, uma concepção de qualidade influenciada pelo modelo gerencial de administração, bastante voltada para os resultados. Nele, a preocupação central da qualidade passa a ser com o controle de indicadores de desempenho das IES e dos cursos, privilegiando como foco os resultados e não os processos educacionais. Em direção contrária, é preciso investir em uma concepção de qualidade que avance “para além da simples padronização de indicadores, abarcando estudos qualitativos e quantitativos refletores da complexidade [da IES]” (MOROSINI, 2014, p. 393).

No campo das hipóteses, admitimos que um razoável número de Comissões Externas nem sempre faça uma apreciação mais acurada do trabalho das CPAs na composição de seu relatório de avaliação institucional, o que gera um efeito em cadeia. Ponderamos isso, uma vez que a Secretaria de Regulação e a Supervisão da Educação Superior (Seres) também não irá considerar isso na hora de elaborar seu parecer de conformidade, recomendando ao Conselho Nacional de Educação (CNE) uma deliberação positiva. Por sua vez, o CNE, no caso de considerar a nota conferida à autoavaliação pela Comissão *in loco*, também não lhe atribui a devida relevância em sua decisão final. Registre-se que tanto a Seres quanto o CNE têm o poder de propor “protocolo de compromisso³ para que a IES supere determinadas falhas”, com base no art. 5º, § 2º, inciso VII e art. 6º, inciso III, do Decreto nº 5.773/2006 (BRASIL, 2006). No caso, a suposição aqui levantada é que este, não obstante ser um “critério de avaliação”, não revela evidências muito claras de que também chega a ser um “critério de regulação”.

3 "Art. 5º, § 2º,- À Secretaria de Educação Superior (leia-se Seres) compete especialmente: [...] VII – celebrar *protocolos de compromisso*, na forma dos arts. 60 e 61; e art. 6º no que diz respeito à matéria objeto deste Decreto"; compete ao CNE: "[...] III – recomendar, por sua Câmara de Educação Superior, providências das Secretarias, entre as quais a *celebração de protocolo de compromisso, quando não satisfeito o padrão de qualidade específico* para credenciamento e credenciamento de universidades, centros universitários e faculdades" (BRASIL, 2006, grifo nosso).

Neste ponto da discussão, é importante proceder à distinção entre os dois critérios mencionados, considerando a política do Sinaes. Com efeito, tal distinção reside no território não muito claro das relações entre avaliação e regulação no Sistema Federal de Ensino (SFE). Após aprovação da Lei nº 10.861/2004, havia a expectativa, por parte dos membros do CNE, de que bastaria a IES obter Conceito “3” na avaliação externa para merecer do Ministério da Educação a homologação da deliberação positiva do referido Conselho. Todavia, na prática, entre abril de 2004 e maio de 2006, quando foi aprovado o Decreto nº 5.773/2006 (BRASIL, 2006), que ficou mais conhecido como “Decreto Ponte”, e transferiu parte das competências do CNE à SESu/MEC, houve disputa de interpretação entre a presidência desse Conselho e a Consultoria Jurídica junto ao Ministério da Educação/Conjur/MEC, a propósito do teor do § único do art. 2º da Lei nº 10.861/2004, na parte em que diz que a avaliação é “referencial básico” dos processos regulatórios. Entendendo que aquilo que é referencial não pode ser considerado determinante, ambos os lados defendiam, quando a situação interessava, ora que a avaliação era determinante, ora que era apenas referencial e, nesta qualidade, dispensável à tomada de decisão.

Do ponto de vista da legislação educacional e da atuação governamental, os processos de avaliação e de regulação são compreendidos como dotados de uma relação de dependência entre si. Isto porque tanto cabe ao poder público fazer a regulação, quanto, igualmente, conduzir os processos de avaliação. Essa relação é tratada no referido decreto, que explicita os procedimentos visando atender as funções de regulação, de supervisão e de avaliação das IES, bem como dos cursos de graduação por elas ofertados. De acordo com o disposto no art. 1º do Decreto nº 5.773/2006 (BRASIL 2006):

§ 1º A regulação será realizada por meio de atos administrativos autorizativos do funcionamento de instituições de educação superior e de cursos de graduação e sequenciais.

§ 2º A supervisão será realizada a fim de zelar pela conformidade da oferta de educação superior no sistema federal de ensino com a legislação aplicável.

§ 3º A avaliação realizada pelo Sistema Nacional de Avaliação da Educação Superior – SINAES constituirá referencial básico para os processos

de regulação e supervisão da educação superior, a fim de promover a melhoria de sua qualidade.

O cenário, antes comentado, revela que se tem intensificado a preocupação entre regulação e avaliação no âmbito do Sinaes, ainda que ambos os processos sejam contemplados na formulação da política. Isto nos leva a compartilhar da preocupação de Dias Sobrinho (2010) de que o “grande desafio é encontrar o ponto de equilíbrio entre a avaliação formativa e a regulação controladora. Isso só pode ser construído por meio do diálogo e do sentido público da responsabilidade social” (p. 223).

Por fim, ressaltamos que consideramos a legitimidade da tese de que nenhuma avaliação é perfeita, visto que, em sua essência, sempre existirão contradições, dilemas e impasses. Todavia, mesmo validando essa premissa, chamamos a atenção para o fato de que os índices, agregados ao Sinaes, como, por exemplo, o Conceito Preliminar de Curso (CPC) e o Índice Geral de Cursos (IGC), ambos criados em 2008, têm ganhado centralidade na política avaliativa aqui analisada. Em contrapartida, a autoavaliação, definida na lei que instituiu o Sinaes como a força motriz a partir da qual os outros processos avaliativos ganhariam existência, tem ficado, em certo sentido, subsumida no processo como um todo.

Conclusão

A análise realizada ao longo deste artigo suscita algumas ponderações, tendo por referência a necessidade da construção de um percurso avaliativo de natureza formativa subjacente à política do Sinaes. Entre essas ponderações, retomamos aqui aquela que se refere à concepção dessa avaliação devido à sua grande importância para que cada IES dela se empodere, visando fortalecer sua própria autonomia.

Sem dúvida, uma avaliação assume um caráter formativo na medida em que se mostra capaz de contribuir, sobremaneira, para a emancipação dos indivíduos. Assim, é preciso que o processo avaliativo, como um todo, seja concebido como algo que propicia

[...] o conhecimento, a interpretação, a tomada de decisões e a transformação da realidade. Não se trata de uma fotografia pontual e estática, que apenas estaria fornecendo elementos fragmentados de conhecimento. Sendo pedagógica, social, organizada e intencionalmente orientada para a transformação, ela mesma é um mecanismo de construção e elevação da qualidade. Os ajustes podem ser feitos ao longo do processo (DIAS SOBRINHO, 2000, p. 195).

Essa abordagem conceitual é especialmente significativa para o contexto da autoavaliação, foco do presente artigo. Entre outras razões, ela é importante para lembrar que, ao produzirem seus relatórios à luz das dimensões apontadas pelos Sinaes – ou de outras que se mostrem organicamente articuladas entre si e com sua missão –, as IES podem avançar em práticas de empoderamento do processo avaliativo. Sem dúvida, nisto reside a possibilidade de cada instituição vir a fomentar a cultura da avaliação na perspectiva formativa, poder criar e/ou fortalecer suas referências, bem como fazer comparações entre resultados obtidos em sua autoavaliação.

Nessa lógica, a autoavaliação pode se configurar como instrumento de gestão pedagógica, considerando o uso dos resultados em relação ao trabalho que a IES realiza. Além disso, assume o papel de fornecer elementos que provoquem um olhar da instituição para dentro de si, na perspectiva de encontrar alternativas que contribuam para melhorar a qualidade da educação que ela oferece.

Em consequência, é de grande importância que as IES avancem na produção dos seus relatórios de CPA. Na prática, é urgente e necessário que, na sistematização desses documentos, elas empreendam esforço visando explicitar os indicadores que norteiam os processos avaliativos realizados em suas práticas cotidianas. Entre outras razões, esse esforço é exigido para promover a devida articulação entre a concepção de avaliação adotada e aquilo que o relatório expressa sobre a realidade institucional.

Para tanto, é preciso que as dimensões e seus indicadores de análise mostrem a devida contextualização, em relação às peculiaridades da IES, para que os resultados da autoavaliação possam assumir rumos que mantenham proximidade com os esperados. Também é importante lembrar que se apoderar da avaliação e satisfazer adequadamente aos seus critérios é uma

iniciativa saudável, senão desejável. Entretanto, quando ocorre o contrário – aqui entendido como renúncia às dimensões do Sinaes e a não proposição de uma metodologia alternativa de autoavaliação pela IES –, as possibilidades de aperfeiçoamento do seu trabalho ficam comprometidas, quando não anuladas. Na prática, este fato torna possível supor que, na proporção em que a instituição não consolida sua autoavaliação, ela poderá revelar dificuldades práticas de diversas naturezas, como, por exemplo, definir, de forma estratégica, suas ações futuras.

As reflexões, propostas ao longo deste artigo, permitem concluir, também, que se torna necessário, cada vez mais, que os relatórios, produzidos pelas CPAs, avancem rumo à intencionalidade explicitada nos marcos do Sinaes de que a avaliação seja emancipatória, considerando o protagonismo da comunidade acadêmica, representada pelos seus vários segmentos. Sob este ângulo, é importante avançar para uma concepção de autoavaliação que, sendo sistemática e processual, apresente possibilidades e contribuições para o aperfeiçoamento da IES. Tais contribuições precisam assegurar, ao mesmo tempo, dois princípios fundamentais no trabalho de cada instituição: uma avaliação formativa e engajada com a transformação social e sua responsabilidade social.

Convém lembrar que, caso a autoavaliação seja esvaziada de seu caráter formativo, não apenas seu caráter emancipatório, mas também as CPAs correm o risco de ter ameaçada sua sobrevivência no contexto mais amplo das IES, ainda que continuem a existir para atender a existências legais. Para Leite (2008), esse risco existe, à medida que outro componente do Sinaes, o Enade, na condição de um instrumento secundário nesse mesmo sistema, tem passado a ganhar destaque e força política.

Por último e considerando o foco adotado neste artigo, parece razoável supor que, no âmbito do Sinaes, a não centralidade da autoavaliação institucional tenha causa, também, na ausência do próprio Estado no que tange ao acompanhamento sistemático e cíclico do trabalho das CPAs. Afinal, o que tem se verificado, na prática, é a postagem dos relatórios de autoavaliação pelas IES, para atender ao ciclo avaliativo, porém sem retorno efetivo do Estado em relação a esse componente da política de avaliação da educação superior brasileira, em vigência desde 2004.

Referências

ACIOLI, A. C; OLIVEIRA, M. A. A. *A qualidade do ensino superior no contexto latino-americano: uma contribuição a partir dos textos publicados pela revista Calidad en la Educación*. In: XXII SEMINÁRIO NACIONAL UNIVERSITAS/ BR, 2014, Natal, p. 224-244.

AUGUSTO, R.; BALZAN, N. C. A vez e a voz dos coordenadores das CPAs das IES de Campinas que integram o SINAES. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, SP, v. 12, n. 4, dez, p. 97-622, 2007.

AZEVEDO, J. M. L. *A educação como política pública*. 4. ed. Campinas, SP: Autores Associados, 2008, p. 79.

BARREYRO, G. B. De exames, rankings e mídia. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 13, n. 3, p. 863-868, nov. 2008.

BELLONI, I. Autoavaliação institucional: um instrumento de democratização da educação. *Linhas Críticas*, Brasília, v. 5, n. 9, p. 31-58, jul./dez. 1999.

BERTOLIN, J. C. G. *Avaliação da qualidade do sistema de educação superior brasileiro em tempos de mercantilização – período 1994-2003*. Tese (Doutorado em Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

BRASIL. Lei nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. *Diário Oficial [da] República Federativa do Brasil*. Brasília, 23 dez., Seção I, p. 27.833-27.841.

_____. Lei nº 10.861, de 14 de abril de 2004a. Institui o Sistema Nacional de Avaliação da Educação Superior – Sinaes e dá outras providências. *Diário Oficial [da] República Federativa do Brasil*. Brasília, n. 72, 15 abr., 2004, Seção I, p. 3-4.

_____. Decreto nº 5.773, de 9 de maio de 2006. Dispõe sobre o exercício das funções de regulação, supervisão e avaliação de instituições de educação superior e cursos superiores de graduação e sequenciais no sistema federal de ensino. *Diário Oficial [da] República Federativa do Brasil*. 10 maio, 2006, Brasília, Seção I, p. 6-10.

_____. Projeto de Lei nº 4.372/2012. Cria o Instituto Nacional de Supervisão e Avaliação da Educação Superior – Insaes, e dá outras providências. Disponível em: <<http://www.camara.gov.br/proposicoes>>. Acesso em: 2 jan. 2018.

_____. Ministério da Educação. Portaria nº 2.051, de 9 de julho de 2004b. Regulamenta os procedimentos de avaliação do Sistema Nacional de Avaliação da Educação Superior (Sinaes), instituído na Lei nº 10.861, de 14 de abril. *Diário Oficial [da] República Federativa do Brasil*, Brasília, n. 132, 17 jul., 2004, Seção I, p. 12.

_____. Ministério da Educação. Portaria Normativa nº 40, de 12 de dezembro de 2007. Institui o e-MEC, sistema eletrônico de fluxo de trabalho e gerenciamento de informações relativas aos processos de regulação, avaliação e supervisão da educação superior no sistema federal de educação, e o Cadastro e-MEC de Instituições e Cursos Superiores e consolida disposições sobre indicadores de qualidade, banco de avaliadores (BASIS) e o Exame Nacional de Desempenho de Estudantes (Enade) e outras disposições. *Diário Oficial [da] República Federativa do Brasil*. Brasília, 29 dez. 2010, n. 249, Seção 1, p. 23-31 [Republicada].

_____. Ministério da Educação. Secretaria de Educação Superior. Portaria SESU/MEC nº 11, de 28 de abril de 2003. Com a finalidade de analisar, oferecer subsídios, fazer recomendações, propor critérios e estratégias para a reformulação dos processos e políticas de avaliação da Educação Superior e elaborar a revisão crítica dos seus instrumentos, metodologias e critérios utilizados. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 30 abr. 2003. Seção 2, p. 19.

_____. Ministério da Educação. Secretaria de Educação Superior. Portaria SESU/MEC nº 19, de 27 de maio de 2003. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, maio 2003.

_____. Ministério da Educação. Portaria nº 92, de 31 de janeiro de 2014. Aprova, em extrato, os indicadores do Instrumento de Avaliação Institucional Externa para os atos de credenciamento, recredenciamento e transformação de organização acadêmica, modalidade presencial, do Sistema Nacional de Avaliação da Educação Superior – SINAES. *Diário Oficial [da] República Federativa do Brasil*, Brasília, DF, 4 fev. 2014, seção I, p. 5.

DIAS SOBRINHO, J. *Avaliação da educação superior*. São Paulo: Cortez, 2000.

_____. Avaliação educativa: produção de sentidos com valor de formação. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 13, n. 1, p. 193-207, mar. 2008.

_____. Avaliação e transformações da educação superior brasileira (1995-2009): do Provão ao Sinaes. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 15, n. 1, p. 195-224, mar. 2010.

GRIBOSKI, C. M.; FERNANDES, I. R. Avaliação da Educação Superior: como avançar sem desqualificar. In: MANCIBO, D.; BITTAR, M.; CHAVES, V. L. J. (Orgs.). *Educação Superior: expansão e reformas educativas*. Maringá, PR: Eduem, 2012, p. 99-126.

INEP. *Sinaes: Sistema Nacional de Avaliação da Educação Superior*. 3. ed. ampliada. Brasília: Inep, 2009.

LEITE, D. Ameaças pós-rankings sobrevivência das CPAs e da autoavaliação. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 13, n. 3, p. 833-840, nov. 2008.

MARX, Karl; ENGELS, Friedrich. *Obras escolhidas em três volumes*. Volume 3. Rio de Janeiro: Vitória, 1963.

MOROSINI, M. C. Qualidade da educação superior e contextos emergentes. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 19, n. 2, p. 385-405, jul. 2014.

PEIXOTO, Maria do Carmo de Lacerda. Avaliação institucional externa no SINAES: considerações sobre a prática recente. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas. 2011, v. 16, n. 1, p. 11-36.

POLIDORI, M. M.; ARAÚJO, C. M. M.; BARREYRO, G. B. Sinaes: perspectivas e desafios na avaliação da educação superior brasileira. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 14, n. 53, p. 425-436, out./dez. 2006.

SOUSA, J. V. Qualidade na educação superior: lugar e sentido na relação público-privado. *Caderno Cedes*, Campinas, v. 29, n. 78, p. 242-256.

SOUSA, J. V. Formação de professores e avaliação da qualidade da educação superior: o curso de Pedagogia no ciclo do Sinaes. In: SOUZA, R. C. C. R.; MAGALHÃES, S. M. O. (Org.). *Formação, profissionalização e trabalho docente: em defesa da qualidade social da educação*. Campinas, SP: Mercado de Letras, 2016. p. 41-66.

SOUSA, J. V. FERNANDES, I. R.; Sinaes ou Neo-Sinaes? A avaliação da educação superior no Brasil à luz do Enade ampliado. In: CUNHA, C.; SOUSA, J. V.; SILVA, M. A. *Expansão e avaliação da educação superior: diferentes cenários e vozes*. Belo Horizonte, MG; Fino Traço; Brasília, DF: Faculdade de Educação da Universidade de Brasília, 2016. p. 191-215.

TAVARES, M. G. M.; PEIXOTO, A. L. V. A.; CORREIA, C. S. Avaliação e qualidade da Educação Superior sob a ótica da cultura da performatividade: uma análise da produção da revista Ensaio no período pós-Sinaes – 2004-2012. In: XXII SEMINÁRIO NACIONAL UNIVERSITAS/BR, 2014. Natal, p. 347-369, 2014.

José Vieira de Sousa

Doutor em Sociologia pela Universidade de Brasília.

Professor na Universidade de Brasília, Brasília (DF), Brasil;

sovieira1@gmail.com

Artigo submetido em 04/01/2018

Aprovado em 25/09/2018

DINÂMICAS DE EVASÃO NA EDUCAÇÃO SUPERIOR BRASILEIRA

DYNAMICS OF DROP-OUT IN BRAZILIAN HIGHER EDUCATION

DINÁMICAS DE EVASIÓN EN LA FORMACIÓN SUPERIOR BRASILEÑA

Luciana Guedes da Silva

Cecília Brito Alves

Joaquim José Soares Neto

Silene de Paulino Lozzi

RESUMO

Este artigo é resultado da dissertação que teve origem na seguinte pergunta: É possível antecipar os índices de evasão na educação superior, prevenindo, assim, o sistema das consequentes perdas sociais e econômicas envolvidas na alta evasão? A pesquisa tem como objetivo geral investigar as características organizacionais e suas influências sobre as taxas de evasão, de forma a antecipar seus efeitos negativos e possibilitar maior preparo por parte dos agentes envolvidos: instituições, Estado e sociedade. Os altos índices de evasão observados, mesmo quando transpostas as dificuldades para a entrada na graduação, revelam a necessidade de se pensar em políticas públicas que visem minimizar as diversas perdas envolvidas no abandono acadêmico. O modelo aqui investigado é baseado em variáveis do Censo da Educação Superior/Inep e demonstrou explicar mais de 10% das taxas de evasão nos cursos de graduação presencial brasileiros.

Palavras-chave: evasão; educação superior; políticas públicas educacionais; regressão linear.

ABSTRACT

This research has origin from the following question: Is it possible to anticipate the rates of evasion in higher education, thus preventing the system of the consequent social and economic losses involved in high evasion? This prognosis describes the original problem of the present research, whose general objective is to investigate

the organizational characteristics and their influences and behaviors on the evasion rates, in order to anticipate their negative effects and to allow a greater preparation on the part of the involved agents: institutions, State and society. The high evasion rates observed, even after transposing the difficulties for graduation, reveal the need to think about public policies that object to minimize the various losses involved in academic abandonment. This model is based on variables the Census of Higher Education/Inep and has explained more than 10% of the dropout rates in Brazilian undergraduate courses.

Keywords: drop out; higher education; educational public policies; linear regression.

RESUMEN

Esta investigación tiene su origen en la siguiente pregunta: ¿Es posible anticipar las tasas de evasión en la educación superior, evitando así el sistema de las consiguientes pérdidas sociales y económicas involucradas en la alta evasión? Este pronóstico describe el problema original de la presente investigación, cuyo objetivo general es investigar las características organizacionales y sus influencias y comportamientos sobre las tasas de evasión, para anticipar sus efectos negativos y permitir una mayor preparación por parte de los agentes involucrados: instituciones, Estado y sociedad. Las altas tasas de evasión observadas, incluso después de transponer las dificultades para la graduación, revelan la necesidad de pensar en políticas públicas que objeten minimizar las diversas pérdidas involucradas en el abandono académico. Se ha demostrado que el modelo, que considera variables encontradas en el Censo de la Educación Superior/Inep, explica más del 10% de las tasas de deserción en los cursos brasileños de pregrado.

Palabras clave: evasión escolar; formación superior; políticas públicas educativas; regresión lineal.

Introdução

A recente ascensão dos países chamados “emergentes” intensificou as discussões acerca das perspectivas de desenvolvimento e seus projetos, no intuito de voltar os investimentos aos aspectos que trariam melhor posição econômica e social a tais Estados no cenário globalizado. Os países que se tornaram referência mundial de riqueza guardam em comum a característica

de terem investido na melhoria da escolaridade de sua população (DIAS SOBRINHO, 2005), aumentando as preocupações sobre o tema em nações que buscam alcançar índices semelhantes de desenvolvimento.

Estudos coordenados pelo Instituto de Pesquisa Econômica Aplicada (Ipea) vêm, repetidamente, alertando para o caráter seletivo e excludente da educação brasileira (CORBUCCI, 2007 e 2014), que decorreria basicamente da influência dos fatores socioeconômicos em dois momentos da trajetória acadêmica do indivíduo: primeiro com a evasão na educação básica e, posteriormente, com o desencorajamento do ingresso na graduação, que é reforçado pelo modelo privatista de expansão da educação superior (PINTO, 2004, BAGGI; LOPES, 2011). Acrescenta-se, ainda, um terceiro momento, que se dá com a desistência da continuidade na educação superior.

Os altos índices de evasão, que vêm sendo observados mesmo quando transpostas as dificuldades para a entrada na graduação, revelam a necessidade de pensar também em políticas públicas que visem minimizar as diversas perdas envolvidas, tais como perdas no nível pessoal, social e financeiro. Ao se explorar as características de evadidos da educação superior, percebe-se que algumas classes populacionais são recorrentes nesse universo e crescem as suspeitas de que alguns grupos têm a motivação mais facilmente afetada pelas dinâmicas universitárias, logo, uma maior tendência à evasão.

A necessidade de se garantir a permanência dos estudantes nos cursos de graduação, para além do acesso, levou os gestores educacionais a direcionarem estratégias de investimento. De acordo com dados do Ministério da Educação (MEC), o investimento em Instituições Federais de Ensino Superior ultrapassou os 25 bilhões de reais em 2010 (BRASIL, 2015). Nesse mesmo ano, o Programa Nacional de Assistência Estudantil recebeu um investimento de 305 milhões, um aumento de quase 55% em relação ao ano anterior, para que, em 2013, quatro anos depois, a taxa de concluintes tenha tido um aumento de somente 15% em relação a 2010, quando o custo anual de um aluno de educação superior era de R\$ 14.763,00 (PORTAL BRASIL, 2010). Esses números revelam os dois lados da mesma moeda: na medida em que demonstram o empenho governamental em cumprir com os objetivos para a educação superior, também revelam perdas financeiras que, quando ocorrem, são substancialmente onerosas ao Estado, especialmente em situações de instabilidade econômica e escassez de

recursos. A evasão dos estudantes constitui uma das causas de perda dos investimentos públicos financeiros.

A evasão da educação superior tem suas consequências em diversos níveis, seja no âmbito dos alunos, das Instituições de Ensino Superior (IES) e, seja no âmbito da sociedade. Embora para os alunos a mudança de curso, que constitui evasão do mesmo, possa se configurar como positiva e possa se refletir em benefícios pessoais, para a IES e até para a sociedade, em termos de investimentos, as consequências são igualmente prejudiciais – seja ela pública ou privada, uma vez que a vaga ocupada não terá como desfecho um concluinte daquele curso. É sob essa perspectiva que este estudo se coloca, investigando as propriedades abarcadas pelas IES a partir do conceito de evasão de curso, isto é, o encerramento do vínculo do aluno com a unidade que realizou sua matrícula no primeiro período examinado.

Buscando entender o fenômeno de evasão, vários autores têm mapeado os fatores associados a esta e elaborado modelos que possam explicá-la. Por exemplo, Page (1998, 2007) relata alguns aspectos que influenciam o abandono do Ensino Superior. Em particular, ele discute quatro categorias de abandono. Segundo o autor, a primeira e maior categoria é o abandono precoce – estudantes se matriculam no curso errado e, assim que eles percebem seu erro, evadem. A segunda categoria é marcada pelo desistente oportunista, que entra na faculdade porque não tem nada mais interessante a fazer e, quando surge uma oportunidade melhor, como se fosse um emprego, ele abandona o curso. Na terceira categoria de abandono, o aluno comporta-se como um consumidor, tão logo está satisfeito com a aprendizagem adquirida, ele sai do curso, estes tendem a valorizar mais a competência à certificação. Na quarta categoria, fazem parte as pessoas que abandonam o curso por conta de crises na vida, por exemplo, acidentes, doenças, infortúnios e crises financeiras.

Um importante modelo, o Modelo de Integração do Estudante (TIE), elaborado por Tinto (1975), busca descrever o caminho que leva à evasão escolar. Tinto ressalta, em seu artigo, que a aplicação da teoria do suicídio ao fenômeno da evasão escolar não rende, por si só, uma teoria de evasão que explique, no nível do indivíduo, os comportamentos de evasão. Em vez disso, seu modelo descreve condições propícias ao encorajamento desse comportamento. Tal modelo é construído sobre dois alicerces básicos: a Teoria do Suicídio, de

Durkheim, e a análise de custo-benefício de escolhas individuais a respeito dos investimentos educacionais.

Segundo o autor, o indivíduo abandonaria a universidade ao perceber uma alternativa para seus investimentos de tempo, energia e recursos que lhe trará maiores benefícios em comparação aos custos de permanência na graduação. É, então, a combinação desses dois elementos que culminaria na evasão.

Apesar de os construtos sobre a instituição estarem presentes no modelo, Tinto não aprofunda suas questões, focando principalmente na integração do aluno na universidade. Portanto, este modelo supõe que a evasão estudantil se dá por meio de uma espécie de desequilíbrio entre a integração acadêmica e a integração social, sendo afetadas ainda pelas metas e compromettimentos (valor atribuído a pertencer a determinada universidade) do aluno com seu curso. Somente quando essas estão em consonância, é possível evitar a deserção.

Reconhecendo o grande valor da educação superior como agente de transformação social, a evasão de estudantes torna-se uma questão central nos debates sobre o desenvolvimento do país. Para tanto, faz-se necessário compreender de que forma as IES podem ou não contribuir para a permanência do aluno até a conclusão de seus estudos, ou seja, quais são as principais características e políticas institucionais de uma IES que podem influenciar que o estudante conclua ou não sua graduação. O presente trabalho tem como objetivo explorar características ou fatores institucionais que podem estar relacionados ao abandono dos cursos de graduação, de forma a antecipar seus efeitos negativos e possibilitar maior preparo por parte dos agentes envolvidos – instituições, Estado e sociedade.

Metodologia

Para verificar a influência das variáveis das IES nas taxas de evasão, são utilizados os dados do Censo da Educação Superior, realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Neste estudo, optou-se por utilizar dados de 2013 e 2014 que, combinados, permitiram verificar o número de matrículas nos cursos de um ano

para o outro. Dessa forma, será possível verificar se, e em que dimensão, as características das IES podem refletir nos graus de desistência do estudante.

Em primeiro lugar, foi utilizado o índice de evasão calculado por Lozzi et al. (2018), e excluídos aqueles que resultaram em números negativos. Os valores negativos são decorrentes de incoerências na coleta dos dados, como, por exemplo, a inexistência de um ou mais elementos da fórmula do índice. Para serem mantidos no banco de dados, os cursos deveriam estar “em atividade” tanto em 2013 quanto em 2014, ter todos os elementos da fórmula com valores válidos, sendo que a quantidade de concluintes, em 2013, deveria ser menor do que a quantidade de matriculados no mesmo ano (denominador positivo), e o número de matrículas também deveria ser maior que o de ingressantes para os dois anos. Essa etapa resultou em um banco de dados com 19.266 cursos.

O conceito de evasão abordado anteriormente dá origem e é representado pela equação:

$$E = (1 - ((M_{2014} - Ig_{2014}) / (M_{2013} - Cc_{2013}))) \times 100$$

em que: M_{2014} = matrículas em 2014;
 M_{2013} = matrículas em 2013;
 Cc_{2013} = concluintes em 2013; e
 Ig_{2014} = ingressantes em 2014.

No segundo momento, excluíram-se os 487 cursos com modalidade de ensino a distância (EaD), resultando em um total de 18.779 cursos presenciais, pertencentes a 1.968 IES. A decisão pela exclusão se deu principalmente devido às especificidades da modalidade EaD. Tanto o modelo teórico utilizado majoritariamente, de Tinto (1973), quanto os que dão suporte adicional (BEAN; METZNER, 1985) tem como um dos pilares básicos para o resultado de evasão a interação do estudante com a instituição e com os pares, o que, para uma modelagem que contemple os cursos EaD, exigiria um estudo específico de sua forma, fugindo ao escopo desta pesquisa. Ainda, o modelo verifica a influência de fatores locais (localização e região), que não fazem sentido quando não há utilização do espaço físico.

Para tornar o estudo o mais abrangente possível, foram mantidos todos os cursos, sendo que as variáveis que integraram a base de dados utilizada são descritas conforme tabela 1.

Tabela 1 Variáveis inseridas no modelo

VARIÁVEL	DESCRIÇÃO DA VARIÁVEL	BANCO DE ORIGEM	TIPO DE VARIÁVEL
1 Feminino	% de mulheres no curso.	Alunos	Contínua
2 Cotas	% de alunos que ingressaram com reserva de vagas.	Alunos	Contínua
3 Apoio Social	% de alunos beneficiários de programa de bolsa ou apoio social/de permanência.	Alunos	Contínua
4 Financiamento	% de alunos que realizaram financiamento (somente para o modelo de IES privadas).	Alunos	Contínua
5 Pretos	% de alunos negros.	Alunos	Contínua
6 Ingresso	% de alunos que ingressaram por outras formas que não o vestibular/Enem.	Alunos	Contínua
7 Região	Região brasileira.	IES	Nominal
8 Categoria Administrativa	Informa se a IES é pública ou privada e qual a sua esfera administrativa.	IES	Nominal
9 Organização Acadêmica	Tipo de IES.	IES	Nominal

VARIÁVEL	DESCRIÇÃO DA VARIÁVEL	BANCO DE ORIGEM	TIPO DE VARIÁVEL
10 Evasão	Índice calculado por Lozzi (2016) é variável derivada, e assume valores entre 0 e 100.	Cursos	Contínua
11 Localização	Informa se a IES se localiza na capital ou não.	IES	Nominal
12 Área geral	Nome da área geral conforme adaptação da classificação internacional Eurostat/Unesco/OCDE.	Cursos	Nominal

Fonte: elaborado pela autora.

Uma análise descritiva dos principais componentes do modelo se faz necessária para melhorar a compreensão do comportamento das variáveis envolvidas no modelo. Os diagnósticos realizados devem ser baseados em três critérios básicos: clareza, simplicidade e autoexplicação (DIETZ; KALOF, 2014). Os gráficos, os agrupamentos e as medidas devem sintetizar a situação analisada, como uma espécie de retrato ou modelo simplificado da realidade, para que as técnicas seguintes sejam aplicadas de maneira correta, garantindo correspondência com a realidade estudada.

A mensuração condicional da taxa de evasão se faz útil para explicar um evento desconhecido, sobre o qual poucos têm controle ou de complexa abordagem. Nesse caso, técnicas probabilísticas atuam como a investigação de algo que não é entendível ou palpável a partir do que é visível, concreto. Assim, uma regressão linear atua como um delineamento da variável desconhecida (mas observável) “evasão” a partir das variáveis explicativas. O modelo geral de regressão linear é descrito a seguir:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

Em que Y é a taxa de evasão, β_0 é a constante, ou seja, o valor que a evasão assume quando todas as explicativas são mantidas constantes e iguais a 0, β_k ($k = 1, 2, \dots, n$) é o coeficiente de cada variável, n é o número de variáveis explicativas no modelo e ε é o erro ou resíduo, que representa as possíveis

variáveis explicativas que não foram mensuradas ou excluídas do modelo. Ao realizar uma modelagem, espera-se reduzir o erro ε ao mínimo possível, preferencialmente até o imensurável acaso. No caso de políticas públicas, entretanto, isso não é possível, pois realidades complexas envolvem dimensões diversas de difícil mensuração. Portanto, esta medida também pode funcionar como indicador da qualidade dos dados, dado que, quanto maior o ε , menor a representatividade das informações disponíveis.

A variável *Região* foi dicotomizada e decomposta em cinco variáveis, quais sejam: Norte, Nordeste, Sudeste, Sul e Centro-Oeste, onde cada variável assume valor 1 quando a IES pertence àquela *Região* e 0 caso contrário; a mesma decomposição foi realizada para área geral. Para a pesquisa apresentada, tem-se $n = 13$, que corresponde ao número de variáveis no modelo. Para categoria *Administrativa* e *Localização*, também assumem valor 1 as IES públicas e localizadas em área urbana, respectivamente. As demais variáveis no modelo são do tipo contínuas, pois, quando importadas do banco de alunos, foram transformadas em percentual.

Neste trabalho, foi realizada a separação de cursos entre provenientes de IES públicas e privadas, uma vez que, para as primeiras, não será inserida a variável “percentual de alunos que realizaram financiamento estudantil” e, para as segundas, a variável excluída é “percentual de alunos cotistas”, considerando as políticas específicas para cada setor. Dessa forma, os modelos finais de evasão proposto são dados por:

$$\begin{aligned} \text{EvasãoPúb} = & \beta_0 + \beta_1 * \text{CatAdm} + \beta_2 * \text{Localização} + \beta_3 * \text{Norte} + \beta_4 * \text{Nordeste} \\ & + \beta_5 * \text{Sudeste} + \beta_6 * \text{Sul} + \beta_7 * \text{CentroOeste} + \beta_8 * \text{PctMulheres} \\ & + \beta_9 * \text{Pct Pretos} + \beta_{10} * \text{Pct ApoioSocial} + \beta_{11} * \text{Pct Ingresso} + \beta_{12} \\ & * \text{Pct Cotistas} + \beta_{13} * \text{Educação} + \beta_{14} * \text{HumArt} + \beta_{15} * \text{CSNDir} + \beta_{16} \\ & * \text{CMC} + \beta_{17} * \text{EngeConstru} + \beta_{18} * \text{AgriVet} + \beta_{19} * \text{SaúdeBES} + \beta_{20} \\ & * \text{Serviços} + \beta_{21} * \text{Medicina} + \varepsilon^1 \end{aligned}$$

e

1 As variáveis foram abreviadas para facilitar a visualização.

$$\begin{aligned}
 \text{EvasãoPriv} = & \beta_0 + \beta_1 * \text{CatAdm} + \beta_2 * \text{Localização} + \beta_3 * \text{Norte} + \beta_4 * \text{Nordeste} \\
 & + \beta_5 * \text{Sudeste} + \beta_6 * \text{Sul} + \beta_7 * \text{CentroOeste} + \beta_8 * \text{PctMulheres} \\
 & + \beta_9 * \text{Pct Pretos} + \beta_{10} * \text{Pct ApoioSocial} + \beta_{11} * \text{Pct Ingresso} + \beta_{12} \\
 & * \text{Pct Financiamento} + \beta_{13} * \text{Educação} + \beta_{14} * \text{HumArt} + \beta_{15} * \\
 & \text{CSNDir} + \beta_{16} * \text{CMC} + \beta_{17} * \text{EngeConstru} + \beta_{18} * \text{AgriVet} + \beta_{19} * \\
 & \text{SaúdeBES} + \beta_{20} * \text{Serviços} + \beta_{21} * \text{Medicina} + \varepsilon
 \end{aligned}$$

Resultado e discussão

O Censo da Educação Superior no Brasil levantou 710 cursos diferentes em 2014, que podem ser agrupados segundo classificação de áreas gerais segundo a Organização para Cooperação e Desenvolvimento Econômico (OCDE) em oito áreas gerais: (i) Educação; (ii) Humanidades e Artes; (iii) Ciências Sociais, Negócios e Direito; (iv) Ciências, Matemática e Computação; (v) Engenharia, Produção e Construção; (vi) Agricultura e Veterinária; (vii) Saúde e Bem-Estar Social; e (viii) Serviços. Com o intuito de evitar variabilidades muito pequenas, bem como evitar resultados demasiadamente pulverizados que dificultem a leitura dos dados, mas levando em conta a relevância do enquadramento de áreas para a investigação, o presente estudo trata-se essencialmente do nível das cinco regiões administrativas. Considerando o enfoque nas taxas de evasão deste trabalho, somente a área de Medicina foi desagregada da área “Saúde e Bem-Estar Social”, por apresentar valor médio muito menor que o dos demais cursos da mesma área.

Os dados apresentados na tabela 2 mostram que a região Sudeste concentra a maior parte de cursos de todas as áreas, seguida das regiões Sul e Nordeste. As maiores taxas médias de evasão são dos cursos de Serviços, Ciências e Matemática.

Tabela 2 Distribuição de cursos segundo área geral por região e evasão média correspondente (Brasil, 2013)

	EDUCAÇÃO	HUMANIDADES E ARTES	CIÊNCIAS SOCIAIS, NEGÓCIOS E DIREITO	CIÊNCIAS, MATEMÁTICA E COMPUTAÇÃO	ENGENHARIA, PRODUÇÃO E CONSTRUÇÃO	AGRICULTURA E VETERINÁRIA	SAÚDE E BEM ESTAR SOCIAL	SERVIÇOS	MEDICINA
Norte	8,5%	3,6%	4,8%	5,4%	4,6%	9,5%	6,0%	5,2%	10,0%
Nordeste	22,0%	14,4%	16,3%	16,1%	13,8%	19,0%	19,0%	18,8%	19,4%
Sudeste	39,4%	48,4%	49,7%	49,8%	54,8%	33,5%	48,7%	46,3%	44,4%
Sul	18,4%	26,7%	20,1%	19,2%	20,5%	22,5%	17,4%	20,5%	18,8%
Centro-Oeste	11,7%	6,9%	9,2%	9,6%	6,2%	15,6%	8,9%	9,3%	7,5%
Evasão Média	23,98	29,84	27,51	30,25	23,54	19,38	22,6	32,82	5,01

Fonte: elaborado com base nos dados de evasão de Lozzi et al. (2017).

A região Sudeste também apresentou maior percentual médio de alunos com ingresso alternativo ao vestibular. Por outro lado, foram encontradas menores taxas de alunos cotistas e pretos do que nas demais regiões, embora seja a região com a maior quantidade de IES do Brasil, conforme apontado na tabela 3.

Tabela 3 Percentuais médios e quantidade total de IES por região (Brasil, 2013)

REGIÃO	% MÉDIO DE MULHERES	% MÉDIO DE ALUNOS PRETOS	% MÉDIO DE ALUNOS QUE RECEBEM APOIO SOCIAL	% MÉDIO DE ALUNOS QUE REALIZARAM FINANCIAMENTO ESTUDANTIL ¹	% MÉDIO DE ALUNOS COTISTAS ²	% MÉDIO DE ALUNOS QUE INGRESSARAM POR FORMAS ALTERNATIVAS	QUANTIDADE MÉDIA DE ALUNOS POR IES	QUANTIDADE DE IES
Norte	56,4	35,8	6,6	37,9	6,3	17,1	23.376	1.108
Nordeste	56,3	28,6	7,8	41,6	15,2	11,7	17.645	3.309
Sudeste	54,3	14,6	7,1	39,8	13,6	14,3	64.079	8.888
Sul	54,2	5,8	6,7	33,8	18,7	20,7	22.148	3.700
Centro-oeste	56,1	23,9	6,3	41,9	20,8	10,2	31.127	1.774
Brasil	54,9	17,4	7,0	38,3	4,6	14,9	42.121	18.779

Fonte: elaborado pela autora com base nos dados de evasão de Lozzi et al. (2017).

Nota: Somente para cursos de IES privadas.¹ Somente para cursos de IES públicas.²

Em relação às taxas de evasão, o percentual de alunos cotistas apresentou relação inversa (com valor negativo), ou seja, quanto maior a quantidade de alunos cotistas observada no curso, menores as taxas de evasão. Esse comportamento já foi observado por Velloso e Cardoso (2008), que analisaram esse fenômeno como o possível reflexo de um maior empenho nos estudos por parte dos que ingressam pela reserva de vagas e apontaram que, por apresentar menor evasão, esse público não deveria ser prioritário no direcionamento de políticas de apoio a alunos com deficiências acadêmicas. Entretanto, como no estudo citado, não foram abordadas relações de causa e efeito, não há evidências de que a evasão entre esses alunos é menor devido à eficácia desses auxílios – tampouco foi explorada suficientemente a tipificação do apoio.

A interpretação desse fenômeno (menor evasão em cursos com maior quantidade de cotistas) exige cautela, pois, ao verificarmos as regiões onde a proporção média de cotistas é maior, encontramos o Centro-Oeste e o Sul em primeiro e segundo lugar (com 20,8% e 18,7% respectivamente), conforme a tabela 3. A região Centro-Oeste, com a segunda maior evasão média do Brasil (26,81%), possui a maior proporção média de cotistas (20,8%), mas é onde estão somente 12% dos cursos de IES públicas. Por outro lado, a região Sudeste, que apresenta a maior taxa de evasão média em 2013 (28,13),

aparece com somente 13,6% de alunos cotistas, mas concentra quase um terço dos cursos públicos (30,7%).

A região Nordeste possui taxa média de evasão e de proporção de cotistas menor que a do Centro-Oeste (com 22,3% e 15,2%), mas 29,6% dos cursos analisados estão nesta região. Isto deixa claro que as proporções demonstram uma tendência que mascara números absolutos, ou seja, numericamente, o Nordeste teria mais cotistas do que o Centro-Oeste, com menor taxa média de evasão, enquanto o Centro-Oeste teria uma quantidade absoluta menor de alunos cotistas, mas com taxa maior de evasão, contrariando a hipótese de que cursos com maior quantidade de cotistas possuem baixa evasão. No caso do Nordeste e do Centro-Oeste, essas regiões são, ainda, muito distintas quanto aos padrões educacionais, especialmente da educação básica: os resultados estaduais do Programa Internacional de Avaliação de Estudantes (Pisa), 2012, colocam o Distrito Federal em segundo lugar no desempenho, sendo que as últimas três posições do *ranking* são ocupadas por estados nordestinos.

Assim, esses dados indicam que pode haver uma relação circunstancial entre evasão e quantidade de cotistas, exigindo maior atenção em seu detalhamento. Sabendo que esses resultados se refletem na entrada na educação superior e em tempos de fortes críticas a programas sociais de acesso e inclusão, selecionar indivíduos para programas de permanência na educação superior, baseado unicamente na correlação linear de evasão e entrada por cotas (tabela 4), pode ser excludente e contraproducente às discussões de desigualdades sociais.

Tabela 4 Correlação de pearson com a taxa de evasão

VARIÁVEL	EVASÃO	
	PEARSON	SIG.
% de mulheres	-0,03	0,000
% de alunos pretos	-0,013	0,082
% de alunos que recebem apoio social	0,019	0,009
% de alunos que realizaram financiamento estudantil	0,009	0,279
% de alunos cotistas	-0,152	0,000
% de alunos que ingressaram por formas alternativas ao vestibular	0,033	0,000
Quantidade de alunos na IES	0,303	0,000

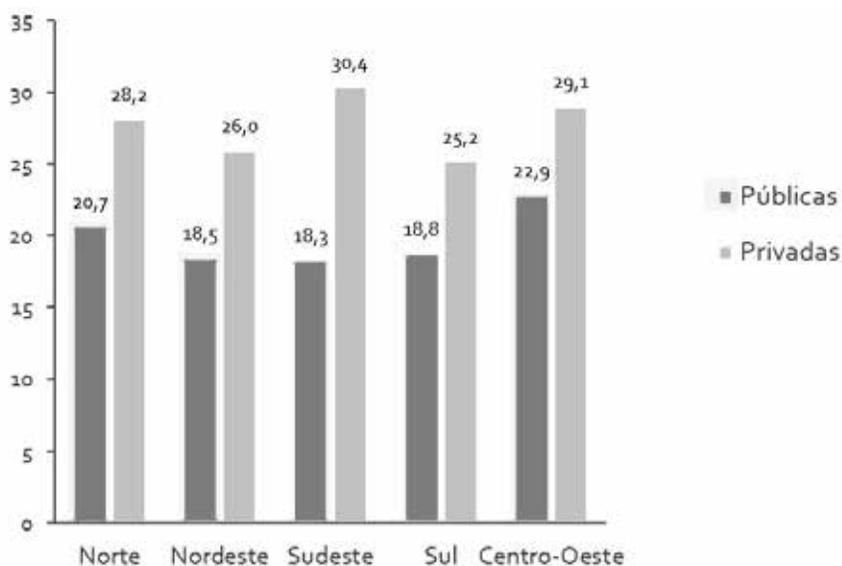
Fonte: elaborado pela autora com base nos dados de evasão de Lozzi et al. (2017).

Ainda foram verificadas diferenças significativas nas taxas de evasão entre as regiões, as áreas gerais e as localizadas em área urbana ou não. As diferenças geográficas e de áreas foram examinadas por meio de um teste de igualdade de médias ANOVA (tabela 5), o qual verifica se existe igualdade entre as médias de evasão para cada uma das categorias de cada variável. Nesse caso, o teste indicou que, ao menos, uma das categorias apresenta taxa de evasão estatisticamente diferente, confirmando as divergências ilustradas no gráfico 1.

Tabela 5 Teste de igualdade de médias anova para evasão entre os grupos de variáveis categóricas (Brasil, 2013)

EVASÃO	F - ENTRE GRUPOS	P-VALOR	ETA ²
Região	98,06	0,000	0,200
Localização	545,18	0,000	0,028
Área Geral	106,64	0,000	0,043

Fonte: elaborado pela autora utilizando dados de evasão de Lozzi et al. (2017).

Gráfico 1 Médias de evasão por região e categoria administrativa

Apesar das particularidades de cada categoria, a começar pela quantidade de cursos públicos e privados existentes (somente 29,2% dos cursos avaliados são de IES públicas), os padrões das taxas de evasão são bastante próximos (tabela 6). A essa regra fogem os cursos das áreas de Saúde e Bem-Estar Social, Engenharia, Produção e Construção e Ciências Sociais, Negócios e Direito, que, no caso das IES públicas, apresentam, em geral, índices de evasão mais baixos do que das IES privadas, com raros casos de taxas maiores do que o intervalo máximo considerado de mais de 70%. Quando combinadas as classes de categoria administrativa e área geral, é possível observar taxas diversas nos cursos da mesma área, caso este seja oferecido por uma IES privada ou pública, sendo que, para todas as áreas, a taxa de evasão em cursos de IES públicas é, em média, menor. Entre os cursos de Medicina, a taxa média de evasão não tem diferença entre IES públicas e privadas, enquanto, entre os demais classificados, como Saúde e Bem-Estar Social, a taxa média de evasão em IES privadas é quase duas vezes maior do que em IES públicas.

Tabela 6 Médias da taxa de evasão para cursos de IES públicas e privadas segundo área geral de enquadramento (Brasil, 2014)

CURSOS	PÚBLICA	PRIVADA
Educação	21,3	26,5
Humanidades e Artes	21,4	34,4
Ciências Sociais, Negócios e Direito	16,0	29,5
Ciências, Matemática e Computação	25,0	33,1
Engenharia, Produção e Construção	16,5	27,4
Agricultura e Veterinária	18,1	21,1
Saúde e Bem-Estar Social	13,2	24,9
Serviços	23,1	35,6
Medicina	5,0	5,0

Fonte: elaborado pela autora utilizando dados de evasão de Lozzi et al. (2017).

Dado que algumas variáveis do estudo só dizem respeito a uma ou outra categoria administrativa – como é o caso de financiamento para IES privadas e cotas para IES públicas, é oportuno que, daqui em diante, os dados sejam analisados separadamente para cada grupo. Para as variáveis categóricas *Região* e *Área Geral do Curso*, foi necessário recodificar as categorias de modo que cada uma delas se tornasse uma variável que recebe valor 1 quando pertence à categoria e 0 caso contrário, como se cada nova variável fosse uma pergunta “*Esse curso pertence à região X?*” em que *X* assume as categorias de região (Norte, Nordeste, Sul, Sudeste e Centro-Oeste). Isto é, um curso da região Nordeste recebe na variável “Nordeste” valor 1, enquanto um curso da região Sul recebe o valor 0 na mesma variável. Já para a variável “Sul”, esses valores invertem-se. Essa mesma lógica é aplicada à *Área Geral do Curso*, totalizando 14 novas variáveis que substituem as duas originais (Educação, Humanidades e Artes, Ciências Sociais, Negócios e Direito, Ciências, Matemática e Computação, Engenharia, Produção e Construção, Agricultura e Veterinária, Saúde e Bem-Estar Social, Serviços e Medicina; e Norte, Nordeste, Sul, Sudeste e Centro-Oeste).

Ao dicotomizar essas duas variáveis, foi necessário definir as categorias que seriam a referência para os modelos. A regressão logística tem sua interpretação baseada em um valor de referência, aquele que o modelo retorna caso todas as variáveis inseridas se mantenham nulas ou controladas.

Assim, a variável com k categorias é substituída por $k-1$ variáveis, uma vez que, quando todas as $k-1$ estão zeradas, necessariamente a k -ésima categoria precisa assumir valor 1. Restam, portanto, quatro novas variáveis de região para substituir a variável original com cinco categorias e oito novas variáveis de área para substituir a original com nove áreas.

Para a escolha das variáveis que seriam a referência para o modelo, ou seja, aquelas para os quais os coeficientes das demais estarão sendo reportados, foi considerado primeiramente o propósito de redução das taxas de evasão. Nesse caso, os cursos da área de Medicina são os que possuem as menores médias de evasão (5% tanto para cursos de IES públicas quanto para de IES privadas). No caso de *Região*, destaca-se a região Sudeste, não por índices de evasão substancialmente diferentes das demais, mas pela alta concentração de cursos que possui. Dessa forma, não estarão explicitamente no modelo descritos os coeficientes relativos às categorias Medicina e região Sudeste.

Em análise de regressão, diz-se que os coeficientes são como a contribuição isolada de cada variável para a resposta. Isso quer dizer que os valores de B para características dos alunos, apresentados nas tabelas 5 e 6, indicam a direção e o número de unidades (como codificadas, neste caso são pontos percentuais) de mudança na variável dependente devido a uma variação de uma unidade em cada variável independente.

Tabela 5 Coeficientes do modelo de cursos de IES públicas (Brasil, 2013)

VARIÁVEL	COEFICIENTES NÃO PADRONIZADOS		COEFICIENTES PADRONIZADOS	T	SIG.	INTERVALO DE CONFIANÇA 95,0% PARA B	
	B	MODELO-PADRÃO				BETA	LIMITE INFERIOR
(Constante)	16,853	2,165		7,783	0	12,608	21,098
Capital	0,743	0,426	0,025	1,743	0,081	-0,092	1,578

VARIÁVEL	COEFICIENTES NÃO PADRONIZADOS		COEFICIENTES PADRONIZADOS BETA	T	SIG.	INTERVALO DE CONFIANÇA 95,0% PARA B	
	B	MODELO-PADRÃO				LIMITE INFERIOR	LIMITE SUPERIOR
% de mulheres	-0,238	0,025	-0,133	-9,529	0	-0,287	-0,189
% de alunos pretos	0,067	0,012	0,086	5,493	0	0,043	0,091
% de alunos que recebem apoio social	0,033	0,011	0,04	2,993	0,003	0,011	0,054
% de alunos que realizaram financiamento estudantil	0,007	0,013	0,007	0,554	0,58	-0,018	0,032
% de alunos que ingressaram por formas alternativas ao vestibular	0,035	0,023	0,022	1,525	0,127	-0,01	0,079
Quantidade de alunos na IES	0	0	-0,103	-7,41	0	0	0
Norte	0,801	0,812	0,016	0,987	0,324	-0,79	2,392
Nordeste	-0,252	0,511	-0,008	-0,493	0,622	-1,254	0,75
Centro-Oeste	4,796	0,645	0,11	7,441	0	3,533	6,06
Sul	1,498	0,552	0,042	2,716	0,007	0,417	2,58
Educação	16,145	1,673	0,545	9,649	0	12,865	19,425
Humanidades e Artes	16,358	1,818	0,262	8,996	0	12,793	19,923
Ciências Sociais, Negócios e Direito	10,658	1,701	0,283	6,267	0	7,324	13,992
Ciências, Matemática e Computação	19,118	1,72	0,454	11,112	0	15,745	22,49
Engenharia, Produção e Construção	9,881	1,723	0,238	5,734	0	6,503	13,259
Agricultura e Veterinária	12,178	1,813	0,201	6,717	0	8,624	15,733

VARIÁVEL	COEFICIENTES NÃO PADRONIZADOS		COEFICIENTES PADRONIZADOS	T	SIG.	INTERVALO DE CONFIANÇA 95,0% PARA B	
	B	MODELO-PADRÃO				LIMITE INFERIOR	LIMITE SUPERIOR
			BETA				
Saúde e Bem-Estar Social	8,522	1,752	0,171	4,865	0	5,088	11,957
Serviços	16,934	2,004	0,186	8,45	0	13,005	20,862

Fonte: elaborado pela autora.

Tabela 6 Coeficientes do modelo de cursos de IES privadas (Brasil, 2013)

VARIÁVEL	COEFICIENTES NÃO PADRONIZADOS		COEFICIENTES PADRONIZADOS	T	SIG.	INTERVALO DE CONFIANÇA 95,0% PARA B	
	B	MODELO-PADRÃO				LIMITE INFERIOR	LIMITE SUPERIOR
			BETA				
(Constante)	5,413	1,887		2,869	0,004	1,715	9,112
Capital	6,367	0,31	0,177	20,514	0	5,759	6,975
% de mulheres	-0,06	0,015	-0,033	-3,976	0	-0,09	-0,031
% de alunos pretos	0	0,009	0	0,031	0,975	-0,018	0,019
% de alunos que recebem apoio social	0,004	0,009	0,004	0,422	0,673	-0,014	0,021
% de alunos que realizaram financiamento estudantil	0,024	0,006	0,035	4,268	0	0,013	0,035
% de alunos que ingressaram por formas alternativas ao vestibular	-0,011	0,007	-0,013	-1,541	0,123	-0,026	0,003

VARIÁVEL	COEFICIENTES NÃO PADRONIZADOS		COEFICIENTES PADRONIZADOS		T	SIG.	INTERVALO DE CONFIANÇA 95,0% PARA B	
	B	MODELO-PADRÃO	BETA	LIMITE INFERIOR			LIMITE SUPERIOR	
Quantidade de alunos na IES	0	0	0,291	34,86	0	0	0	
Norte	-2,147	0,701	-0,026	-3,065	0,002	-3,521	-0,774	
Nordeste	-3,738	0,471	-0,071	-7,94	0	-4,661	-2,815	
Centro-Oeste	-1,393	0,534	-0,022	-2,611	0,009	-2,439	-0,347	
Sul	-2,034	0,386	-0,046	-5,277	0	-2,79	-1,279	
Educação	21,249	1,672	0,431	12,707	0	17,971	24,527	
Humanidades e Artes	26,772	1,774	0,301	15,09	0	23,294	30,25	
Ciências Sociais, Negócios e Direito	23,205	1,649	0,649	14,075	0	19,973	26,437	
Ciências, Matemática e Computação	25,896	1,696	0,435	15,272	0	22,572	29,219	
Engenharia, Produção e Construção	20,322	1,696	0,348	11,984	0	16,998	23,646	
Agricultura e Veterinária	16,978	1,948	0,125	8,715	0	13,159	20,797	
Saúde e Bem Estar Social	18,676	1,673	0,376	11,161	0	15,396	21,956	
Serviços	28,286	1,793	0,296	15,777	0	24,772	31,8	

Fonte: elaborado pela autora utilizando dados de evasão de Lozzi et al. (2017).

Como pode ser observado na tabela 5, os valores de B para *dummies* regionais indicam que, tomando a região Sudeste como referência e mantidas as demais variáveis constantes, os cursos públicos da região Norte têm taxa de evasão cerca de 0,8 ponto percentual (pp) a mais do que os da região Sudeste. Já para os cursos de mesma categoria administrativa, mas localizados na região Centro-Oeste, a taxa de evasão é quase 5 pp maior

do que entre os da região Sudeste e, para os da região Sul, 1,49 pp superior. Os cursos públicos da região Nordeste têm evasão menor do que os do Sudeste em 0,25 pp, no entanto, os resultados do teste *T*, que nos dizem se a evasão de cada uma das regiões é significativamente diferente do Sudeste, mostram que a diferença entre as taxas de evasão dos cursos públicos localizados no Sudeste e no Norte e entre os do Sudeste e do Nordeste é insignificante ($\text{sig.} > 0.05$). No caso dos cursos privados, todas as regiões mostraram-se significativamente diferentes da região Sudeste, segundo o teste *T* (tabela 6). Destaca-se ainda que, para essa categoria administrativa, todas as regiões apresentam percentuais de evasão pelo menos 1 pp menor do que os da região Sudeste, chegando até a uma diminuição de 3,74 pp no Nordeste ($B = -3,74$).

O percentual de mulheres mostrou-se significativo para a evasão, sendo que, nas instituições públicas (ver tabela 5), o aumento de cada ponto percentual na proporção de mulheres reduz a evasão cerca de 0,238 pp, enquanto, naqueles de administração privada, essa redução é somente de 0,06 pp. Outro coeficiente expressivo e significativo, tanto para a categoria pública, quanto para a privada, foi o do curso de Medicina, que nos dois casos é negativo, ou seja, diminui a probabilidade de evasão. Para cursos privados, o coeficiente reduz mais do que em IES públicas conforme o esperado para cursos que exigem grandes investimentos. Os resultados mostram que cursos com maior proporção de mulheres tendem a possuir menores índices de evasão. Quanto aos efeitos de alunos que se autodeclararam pretos, o aumento de 1 pp na proporção não traz praticamente nenhum aumento na evasão em cursos privados (B de aproximadamente 0) e de 0,067 pp para cursos públicos.

As taxas de evasão dentro das áreas gerais dos cursos são significativamente maiores do que as de Medicina, independentemente da categoria administrativa. Entretanto, para os cursos de IES públicas, aqueles da área de Ciências, Matemática e Computação possuem cerca de 19,12 pp a mais de evasão quando comparados com os cursos de Medicina, enquanto os classificados como Saúde e Bem-Estar Social (exceto Medicina) possuem taxas de evasão que se sobressaem em 8,5 pp. As altas taxas de evasão nos cursos de Ciências, Matemática e Computação (habitualmente chamados de Ciências Exatas) nas IES públicas são conhecidas, assim como os altos

índices de retenção. Algumas iniciativas isoladas têm melhorado esse quadro, como a relatada por Machado, Melo Filho e Pinto (2005) para o curso de Química na Universidade Federal do Rio de Janeiro (UFRJ). Esses autores relatam que mudanças simples foram feitas atacando os dois principais problemas responsáveis pela evasão, ou seja, o desconhecimento do curso/carreira e um apoio para os alunos que precisavam entrar mais rapidamente no mercado de trabalho. Embora esses problemas possam ser encontrados em qualquer outro curso, os cursos da área de Ciências Exatas demandam um esforço por vezes inédito, considerando que o brasileiro médio possui desempenho inferior em Matemática em toda a trajetória escolar que antecede o ingresso na educação superior.

Os coeficientes B dos cursos privados (tabela 6) mostram que, mantidas as demais variáveis constantes, a taxa de evasão é maior em quase 30 pp quando o curso pertence à área de Serviços, se comparado com os de Medicina. Todas as áreas apresentaram índices de evasão maiores que os de Medicina. O modelo de cursos de IES privadas apresentou coeficientes de área geral muito superiores àqueles de IES públicas com as maiores diferenças – em relação ao curso de Medicina – encontradas nos cursos das áreas de Serviços ($B = 28,286$), Humanidades e Artes ($B = 26,772$) e Ciências, Matemática e Computação ($B = 25,896$).

A localização na capital mostrou-se uma variável relevante para a explicação da evasão nos cursos privados, mas não para os públicos, apontando que a evasão é maior em cursos nas capitais, em aproximadamente, 6 pp, embora a quantidade absoluta de cursos seja maior fora das capitais, verifica-se que, entre os cursos de IES Públicas, a taxa de evasão não é tão diferente quanto à sua localização, mas, entre os cursos de IES privadas, a evasão é maior nos cursos localizados nas capitais.

O modelo para cursos de IES públicas apresentou R^2 de 0,116 e, para cursos de IES privadas, de 0,18. O coeficiente de determinação R^2 mostra o ajustamento do modelo, pois mede a porcentagem de variação da taxa de evasão explicada pela relação com as variáveis preditoras do modelo. Essa variação é interpretada como o modelo explica a resposta (taxa de Evasão). Assim, podemos dizer que, para cursos de IES públicas,

o modelo explica cerca de 11,6% da evasão, enquanto, para cursos de IES privadas, o modelo testado explica 18%.

Nesse ponto, a relevância das políticas públicas para a Educação Superior brasileira é evidenciada: dado que instituições ligadas à Administração Pública são mais receptivas quanto ao cumprimento das políticas para o setor, evidentemente por seu caráter social, os fatores administrativos influenciam a evasão estudantil em menor intensidade. Por outro lado, o fato de o modelo ter se ajustado melhor para cursos privados deve ser levado em conta, pois há a possibilidade de uma resistência mais forte naquilo que permanece afetando a evasão. Além disso, temos ainda a motivação do estudante na escolha por um curso público ou particular, que geralmente afeta o seu comprometimento e a interação do percurso acadêmico, causando as altas taxas que têm preocupado os gestores.

Considerações finais

Os aspectos aqui abordados são, em grande parte, de difícil atuação, como é o caso das diferenças regionais. Entretanto, conhecer o universo e reiterar essas diferenças reforçam a urgência de se pensar em políticas menos uniformes e mais sensíveis às particularidades de cada cenário. A replicação do estudo em outros períodos, bem como em diferentes recortes e interação com variáveis adicionais, pode auxiliar na construção de indicadores mais consistentes de evasão, além de contribuir para o desenvolvimento de parâmetros acompanháveis mais completos de qualidade da educação superior. O presente trabalho reiterou o apontamento de Baggi e Lopes (2011) de que a investigação mais profunda e efetiva da evasão é tão urgente quanto indissociável de uma avaliação institucional.

Ressalta-se que atualmente os índices médios condizem com aqueles mirados pela Comissão Especial de Estudos sobre Evasão nas Universidades Públicas Brasileiras de 1996 (à época, a Comissão propôs a redução de 50% para 20% nas Instituições Federais de Ensino Superior), mas os custos sociais e financeiros permanecem sendo prejudiciais para a educação, de forma que uma nova reunião, com uma visão atualizada sobre o cenário atual da educação superior brasileiro, seria valiosa.

Dentro da agenda de políticas públicas, o direcionamento de ações acontece conforme a ênfase desejada. Assim, o estudo do cenário de maneira mais abrangente, com maximização das dimensões envolvidas, favorece o entendimento da realidade e identificação dos pontos de melhoria. Ao longo desta pesquisa, foram verificados arranjos e correlações que aparentam ligação com altas taxas de evasão, fazendo com que uma revisão dos fatores associados à evasão, contemplando os diferentes achados em cada curso específico, seja indicada como futura agenda de pesquisa, com o propósito de desvendar os resultados e efeitos esperados de cada variável nas taxas de evasão. Uma revisão dos instrumentos que foram aplicados nos estudos de caso contribuiria para a construção de uma avaliação do Ensino Superior, além de valorizar as pesquisas realizadas no setor.

A especificidade dos cursos e da maneira como a evasão ocorre em cada um deles demonstra que o tema deve ser tratado segundo as particularidades dos cursos e, portanto, a aplicação de uma política administrativa única para as instituições que não leva em consideração as dinâmicas internas aos cursos marginaliza a questão da evasão na educação superior. Estudos em áreas específicas podem gerar indicadores mais confiáveis, mas é necessário que essa procura seja alinhada com a perspectiva do devido tratamento diferenciado dentro das IES.

Ao longo do desenvolvimento desta pesquisa, destaca-se a presença constante de uma carência no que diz respeito a indicadores de “demanda” pelo Ensino Superior. Em todas as etapas, foi verificada uma lacuna de informações que conectem as características de origem e de histórico dos alunos com as dinâmicas dentro da graduação, o que motiva a continuidade de um estudo dedicado a esse tema.

Faz-se necessário reconhecer que o estudo ora apresentado é limitado quanto às suas fontes, que, na ausência de uma avaliação desenvolvida para o propósito de investigação da trajetória acadêmica estudantil, utiliza-se de uma quantidade elevada de variáveis indiretas. Ainda, a utilização de somente um período de transição traz alguma restrição aos resultados apresentados, de modo que o mais indicado seria um acompanhamento longitudinal do fenômeno. Ainda no que diz respeito aos bancos de dados, foi encontrado um número elevado de informações errôneas, imprecisas e incompletas, como, por exemplo, cursos que apresentaram número de

matriculados menor que o de ingressantes. Esse tipo de falha prejudica verificações da realidade como um todo. Essas incorreções são ainda mais visíveis na medida em que a unidade de observação se restringe, com uma larga inconsistência, nas variáveis de *background* do aluno.

São desconhecidos, neste estudo, os efeitos das políticas de transferência entre IES, que poderiam trazer maior elucidação ao entendimento da trajetória do aluno. O acompanhamento dessa trajetória, com inclusão dos dados de entrada (desempenho no vestibular, por exemplo) e atribuição de um número de identificação único que permita a observação longitudinal do estudante no sistema educacional traria ganhos oportunos. Pesquisas qualitativas, com entrevistas com todos os agentes da educação superior (gestores, docentes e alunos), trazem também uma contribuição importante e trariam relevância única ao abordar as definições de qualidade da Educação Superior.

Referências

ALTMANN, H. Influências do Banco Mundial no projeto educacional brasileiro. *Educação e Pesquisa*, São Paulo, 28, p. 77-89, 2002.

BAGGI, Cristiane Aparecida dos Santos; LOPES, Doraci Alves. Evasão e avaliação institucional no ensino superior: uma discussão bibliográfica. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, p. 355-374, jul. 2011. Disponível em: <<http://www.scielo.br>>. Acesso em: 16 maio 2016.

BANCO MUNDIAL. *Aprendizagem para todos. Estratégia 2020 para a Educação do Grupo Banco Mundial*. Resumo Executivo. 2011.

BEAN, John P.; METZNER, B. S. A Conceptual Model of Nontraditional Undergraduate Student Attrition. *Review of Educational Research*, 55(4), 485-540. 1985. Disponível em: <<http://www.jstor.org/stable/1170245>>. Acesso em: 16 maio 2016.

BRASIL. *Lei nº 13.005*, de 25 de junho de 2014. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2014/Lei/L13005.htm>. Acesso em: 20 abr. 2018.

_____. Ministério da Educação. Secretaria de Ensino Superior. *A democratização e expansão da educação superior no país: 2003-2014*. Brasília, 2015. Disponível em: <http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=16762-balanco-social-sesu-2003-2014&Itemid=30192>. Acesso em: 13 fev. 2017.

CORBUCCI, P. R. *Desafios da educação superior e desenvolvimento no Brasil*. Brasília: Ipea, 2007.

_____. *Evolução do acesso de jovens à educação superior no Brasil*. Brasília: Ipea, 2014.

DIAS SOBRINHO, José. Educação superior, globalização e democratização: qual universidade? *Revista Brasileira de Educação*, n. 28, p. 164-173, 2005.

DIETZ, T.; KALOF, L. *Introdução à Estatística Social*. São Paulo: LTC, 2014.

FERNANDEZ LAMARRA, N. Universidad y calidad en América Latina en perspectiva comparada: Interrogantes y desafios. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, Sorocaba, v. 17, n. 3, p. 661-668, nov. 2012. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-40772012000300006&lng=en&nrm=iso>. Acesso em: 13 Feb. 2017.

_____. *O fenômeno da evasão escolar na educação superior no Brasil*. 75 f. Dissertação (Mestrado em Educação) – Programa de Pós-Graduação em Educação da Universidade Católica de Brasília, Brasília, 2005.

GAIOSO, Natália Pacheco de Lacerda. *O fenômeno da evasão escolar na educação superior no Brasil*. Repetencia y Desercion Universitária em América Latina. Santiago, Chile: Unesco, 2006.

INEP. *Indicadores Financeiros Educacionais – Brasil 2000-2014*. Brasília: Inep, 2016. Disponível em: <<http://portal.inep.gov.br/web/guest/indicadores-financeiros-educacionais>>. Acesso em: 13 fev. 2017.

LOZZI, Silene de Paulino. *Evasão nos cursos de Farmácia de instituições do ensino superior públicas e privadas no biênio 2013-2014*. In: 6ª CONFERÊNCIA

DA FORGES, subordinada ao tema: Para que(m) servem a Universidade e as Instituições do Ensino Superior? Balanços, Proposições e Desafios Acerca do Papel das IES no Séc. XXI. No prelo 2016.

LOZZI, Silene de Paulino; ALVES, Cecília Brito; GUEDES, Luciana; SOARES-NETO, Joaquim José *A evasão na educação superior no Brasil*. Avaliação, ISSN 1414-4077. No prelo 2017.

MACHADO, Sérgio P.; MELO FILHO, João Massena; PINTO, Angelo C. A evasão nos cursos de graduação de química: uma experiência de sucesso feita no Instituto de Química da Universidade Federal do Rio de Janeiro para diminuir a evasão. *Química Nova*, v. 28, n. 0, p. 41-43, 2005.

NERI, Marcelo. *Motivos da evasão escolar*. Brasília: Fundação Getúlio Vargas 2009.

PAGE, Mich. Numbers and narratives: what can schoolteachers tell us about college drop-out? *Research in Post-Compulsory Education*, 9:2, 239-248, 2007. DOI: 10.1080/13596740400200177.

_____. *Collecting Slices of College Dropouts' Lives*. In: ERBEN, M. (Ed.). *Biography and Education: a reader*, London: Falmer Press. Social Research and Educational Studies, Series 19, 1998.

PINTO, J. M. R. O acesso à educação superior no Brasil. *Educação & Sociedade*, v. 25, n. 88, p. 727-756, out. 2004. Disponível em: <<http://www.cedes.unicamp.br>>. Acesso em: 13 fev. 2017.

PORTAL BRASIL. *Aluno universitário custa cinco vezes mais que um estudante da educação básica*. 2010. Disponível em: <<http://www.brasil.gov.br/educacao/2010/03/aluno-universitario-custa-cinco-vezes-mais-que-um-estudante-da-educacao-basica>>. Acesso em: 13 fev. 2017.

TINTO, Vincent. Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), p. 89-125, 1975. Disponível em: <<http://www.jstor.org/stable/1170024>>. Acesso em: 16 maio 2016.

VELLOSO, Jacques; CARDOSO, Claudete Batista. *Evasão na Educação superior: alunos cotistas e não-cotistas na Universidade de Brasília*. In: REUNIÃO ANUAL DA ANPED, v. 31, 2008.

Luciana Guedes da Silva

Mestre em Desenvolvimento, Sociedade e Cooperação Internacional pela Universidade de Brasília, Brasília (DF), Brasil;
luciguedes@gmail.com

Cecília Brito Alves

Doutora em Psicologia Educacional pela Universidade de Alberta, Canadá.
Professora da Universidade de Brasília, Brasília (DF), Brasil;
cecilia.brito.alves@gmail.com

Joaquim José Soares Neto

Doutor em Química Teórica. pela Aarhus University, Dinamarca.
Professor da Universidade de Brasília, Brasília (DF), Brasil;
jjsoaresneto@gmail.com

Silene de Paulino Lozzi

Doutora em Ciências Biológicas pela Universidade de Brasília.
Professora da Universidade de Brasília, Brasília (DF), Brasil;
silozzi@gmail.com

A VALIDADE DO ENADE PARA AVALIAÇÃO DE CURSOS SUPERIORES: UMA ABORDAGEM MULTINÍVEL¹

THE VALIDITY OF ENADE TO ASSESS THE QUALITY OF THE
UNDERGRADUATE COURSES: A MULTILEVEL PERSPECTIVE

LA VALIDEZ DEL ENADE PARA LA EVALUACIÓN DE CURSOS SUPERIORES:
UN ENFOQUE MULTINIVEL

Ricardo Primi

Marjorie Cristina Rocha da Silva

Daniel Bartholomeu

RESUMO

O presente estudo se propõe a identificar qual parcela da variância da prova do Enade está associada ao que se pretende medir e, assim, identificar a sua capacidade de se tornar um indicador da qualidade dos cursos. Foram utilizados dados de 71.838 ingressantes e 58.428 concluintes do curso de Administração no Enade de 2006, além de informações sobre as instituições e o perfil socioeconômico dos respondentes. Os resultados demonstram que as notas do Enade decorrem de diferenças importantes de níveis de habilidade no momento de entrada no curso, e que a maior parte da variabilidade das notas do Enade está dentro dos cursos, explicada pelos perfis socioeconômicos.

Palavras-chave: aprendizagem; definição de pontos de corte; qualidade da escola; competências e habilidades.

¹ Este trabalho foi produzido com o financiamento do Observatório da Educação (Capes/Inep) e do CNPq.

ABSTRACT

The present study aims to identify which part of the variance of the Enade test is associated with what it intends to measure, that is, its capacity to become an indicator of the undergraduate course's quality. Data from 71,838 new freshmen and 58,428 undergraduates of the Administration course at Enade 2006 were used, as well as information about the institutions and the socioeconomic profile of the respondents. Results show that the Enade scores is related to differences in skill levels at the entrance, and that most of Enade's variance resides within courses, explained by socioeconomic profiles.

Keywords: learning; definition of cut-off points; quality of school; competencies and abilities.

RESUMEN

El presente estudio se propone a identificar qué parcela de la varianza de la prueba del Enade está asociada a lo que se pretende medir y así su capacidad para convertirse en un indicador de la calidad de los cursos. Se utilizaron datos de 71.838 ingresantes y 58.428 concluyentes del curso de Administración del Enade en 2006, además de informaciones sobre las instituciones de enseñanza y el perfil socioeconómico de los colaboradores. Los resultados exponen que los distintos conceptos implican diferencias importantes de niveles de habilidad en el momento de entrada en el curso. Y que la mayor parte de la variabilidad de las notas del Enade se refiere al período en que los estudiantes están realizando los cursos y se explica por sus perfiles socioeconómicos.

Palabras clave: aprendizaje; definición de puntos de corte; calidade de la escuela; habilidades y competencias.

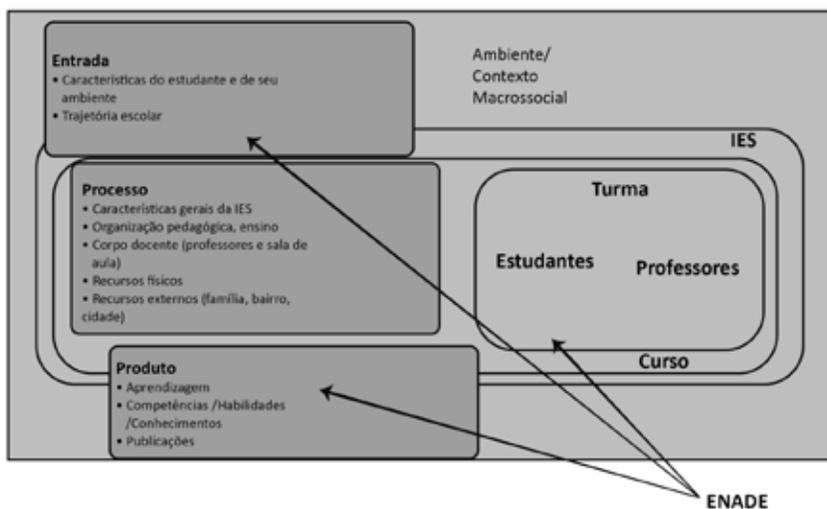
A organização multinível no Enade e sua validade

As avaliações em larga escala, desenvolvidas por entidades públicas, têm um papel fundamental para a sociedade: buscam levantar conhecimentos sobre a eficiência e a qualidade das organizações que provêm bens públicos fundamentais à população, tais como saúde, educação e segurança. Essas

informações são importantes para a gestão eficaz dos recursos públicos, uma vez que clarifica virtudes e falhas do sistema, para que ações interventivas e regulatórias sejam criadas com o objetivo amplo de melhorar sua qualidade, otimizando a utilização dos recursos públicos.

Entre as avaliações adotadas no Brasil, encontra-se o Exame Nacional de Desempenho de Estudantes (Enade). Este é um exame aplicado em larga escala, composto por questões referentes a conteúdos gerais e específicos, elaboradas com o objetivo de aferir as habilidades acadêmicas e as competências profissionais desenvolvidas pelos estudantes dos cursos de graduação das instituições de educação superior (IES). Além disso, busca reunir informações relativas às características socioeconômicas dos estudantes (LIMANA; BRITO, 2005). Embora o Enade avalie o desempenho dos estudantes, sua finalidade é medir a qualidade dos cursos de graduação no que tange a sua contribuição para o desenvolvimento de competências, habilidades e conhecimentos por seus estudantes. A figura 1 ilustra a organização hierárquica alinhada entre IES, cursos, turmas e estudantes, que é inerente a esse sistema de avaliação.

Figura 1 Estrutura multinível inerente ao Enade



Fonte: autoria própria.

Nesse modelo de avaliação há uma suposição implícita de que o curso terá efeito nos alunos ao criar condições favoráveis à aprendizagem e, por meio dela, ao desenvolvimento de conhecimentos, habilidades, competências e atitudes. É importante conceber, de maneira clara, os níveis em que esses dois objetos (estudantes e cursos) da avaliação acontecem. O curso (nível 2), hierarquicamente superior, abriga estudantes (nível 1) de tal forma que qualquer processo pedagógico executado impactará em todos que estejam aninhados nele. Assim, ocorre uma segunda suposição: de que os cursos com diferentes níveis de qualidade terão impacto diferenciado, sendo que cursos de alta qualidade provocarão desenvolvimento em nível mais elevado do que cursos de baixa qualidade. Presume-se uma relação direta entre a qualidade do curso e o desempenho de seus estudantes. Sendo assim, o desempenho agregado dos estudantes de um curso poderia compor um índice ligado à qualidade do curso com base no produto que este gera (aprendizagem). As diferenças entre os índices de distintos cursos demonstrariam as variações de qualidade entre eles. Portanto, avalia-se o desempenho dos estudantes (nível 1) para se obter medidas de qualidade dos cursos (nível 2). Como resultado, esses índices agregados são transformados em cinco notas (1 a 5), consubstanciando os conceitos dos cursos.

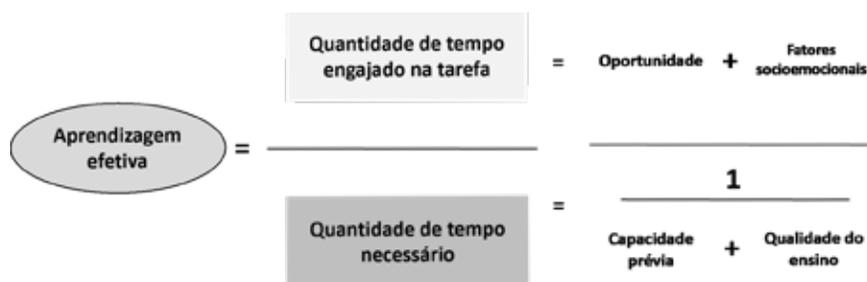
Na literatura psicométrica, essa questão é tratada dentro do conceito de validade. A validade de um sistema de avaliação diz respeito a se ele avalia o que se propõe e quão bem o faz (PRIMI; MUNIZ; NUNES, 2009; ANASTASI; URBINA, 2000; American Educational Research Association – AERA, American Psychological Association – APA, National Council on Measurement in Education – NCME, 1999; URBINA, 2004). Devido ao impacto social que tais sistemas possuem, é urgente que estudos sejam desenvolvidos para certificar a validade desses sistemas de avaliação.

Uma questão central sobre a validade do Enade se refere à possibilidade de se avaliar a qualidade dos cursos por intermédio do desempenho dos estudantes (PRIMI, 2006a, 2006b), já que a literatura sobre o tema discute as evidências do impacto de instituições escolares (IES e cursos) na aprendizagem de seus alunos (por exemplo, SOARES, 2004; 2007). Porém, existem muitos questionamentos sobre se o impacto de cursos com diferenciados níveis de qualidade é substancialmente forte e nítido a tal ponto de ser possível avaliar a qualidade de um curso a partir da aprendizagem dos alunos.

Modelo de Carroll para a aprendizagem escolar

Carroll (1963) concebeu a aprendizagem como resultado da razão entre dois componentes: o tempo despendido no objeto a ser aprendido dividido pelo tempo necessário para aprender aquele objeto (figura 2). O estudante terá sucesso na aprendizagem se dedicar o tempo necessário à tarefa que precisa aprender. Mas há fatores que influenciam no tempo necessário e dedicado à aprendizagem. O tempo necessário depende, por um lado, das capacidades cognitivas como conhecimentos prévios e capacidade de aprender e, por outro, da qualidade do ensino. O ensino é o processo de organização dos materiais e da instrução que exibirá o objeto da aprendizagem em passos sistemáticos com vistas a facilitar a eficácia da aprendizagem. O tempo dedicado depende, por um lado, da oportunidade em si de aprender determinado objeto e, por outro, da motivação (persistência, determinação), entendida como um processo de aprendizagem autorregulada com investimento de tempo e esforço na tarefa.

Figura 2 Modelo de aprendizagem escolar de Carroll (1963)



Fonte: autoria própria.

Nesse modelo, capacidade, conhecimento prévio e fatores socioemocionais (motivação, por exemplo) são aspectos observados no processo de aprendizagem dos alunos (nível 1, destacado em cinza escuro), e oportunidade e qualidade do ensino são aspectos observados no que se refere ao curso (nível 2, destacado em cinza claro). O ponto central aqui é o reconhecimento da multideterminação da aprendizagem (SOARES, 2005; GOLDSTEIN; WOODHOUSE, 2008). Assim, embora os cursos tenham impacto na aprendizagem, o qual decorre do modo como é realizado o contato com o objeto de estudo e da criação de situações eficazes para promover a

aprendizagem – isto é, por meio da qualidade do ensino –, eles não são o único fator. Conhecimento prévio, capacidade e motivação do aluno são fatores bastante importantes que afetam simultaneamente a aprendizagem.

Problematização da validade do Enade

Em síntese, a aprendizagem é medida pelo Enade com base em informações de várias fontes: nível 2 – parte associada ao curso (qualidade e oportunidade); e nível 1 – parte associada aos estudantes e a seus recursos antes de entrarem no curso. Nesse cenário, a questão principal – “É possível avaliar a qualidade do curso a partir do desempenho dos alunos?” – é analisada em dois aspectos: i) é possível identificar e isolar a parcela relevante – de variância confiável de construto relevante –, que contém informações potencialmente ligadas à qualidade do curso; e ii) em caso afirmativo, a quantidade de parcela isolada será suficiente para distinguir cinco grupos de níveis diferenciados de qualidade?

Identificação da parcela relevante

O principal ponto a ser questionado é justamente a utilização do índice de desempenho médio dos concluintes de um curso, comparado relativamente ao índice de outros cursos, como indicador direto de sua qualidade. Essa questão é justificada pela desigualdade do perfil dos ingressantes em diferentes cursos. Dessa forma, as notas dos concluintes “carregam” consigo a influência das diferenças prévias ao efeito que o curso produz no aluno (GOLDSTEIN et al., 1993; SOARES; CASTRO; RIBEIRO, 2000; GOLDESTEIN, 2001; LANDEIRA-FERNANDEZ; PRIMI, 2002; PRIMI; FERNANDEZ; ZIVIANI, 2003; RAUDENBUSH, 2004).

Sabe-se que os cursos com maior demanda passam a selecionar um grupo de estudantes com características diferentes, o que resulta em grupos desiguais entre instituições já no início do percurso acadêmico. Assim, há grupos de alunos com um conjunto de competências, conteúdos e habilidades já adquiridos ao entrar no curso, enquanto há outros com um conjunto mais modesto de recursos. O primeiro grupo iniciará o curso a partir de um ponto mais avançado que o segundo, o que refletirá em sua aprendizagem. Portanto, as diferenças de desempenho (aprendizagem) que se observam ao final do curso não podem ser explicadas exclusivamente

pelo que o curso agregou ao aluno, pois podem refletir e ser confundidas com diferenças preexistentes no perfil do estudante ingressante. Isso tem sido referido como viés de seleção, variáveis omitidas ou endogeneidade na econometria (ANTONAKIS et al., 2010).

Assim, parte das diferenças nos resultados obtidos pelos concluintes, que é explicada por essa diversidade de estudantes no início do curso, não poderá ser atribuída ao que o curso agrega, uma vez que, nesse momento inicial, o aluno ainda não está sob influência do curso. Portanto, quando se observa uma correlação significativa entre o perfil do ingressante e as medidas de qualidade obtidas com os concluintes, já se sabe que uma parcela da variância do desempenho desses últimos não pode ser usada como indicador da qualidade do curso. Desse modo, torna-se, então, necessário aplicar algum procedimento de “correção” da nota do aluno concluinte, na tentativa de expurgar diferenças associadas ao perfil inicial, com vistas a poder produzir indicadores mais depurados daquilo que, por hipótese, poderia estar associado ao efeito do curso (RAUDENBUSH, 2004a, 2004b; RUBIN, 1974).

Uma possibilidade de realizar esse controle é por meio da medida de valor agregado. Assim, como afirma Raudenbush:

[...] um tema comum na literatura é que nós devemos avaliar escolas e professores por intermédio de comparações de seus ‘valores agregados’ à aprendizagem do aluno, ao invés de níveis de desempenho não ajustados ou, como na prática corrente, a proporção de alunos em uma escola ou classe identificada como proficientes (2004b, p. 121).

O valor agregado é definido como a taxa de mudança ocorrida em um determinado tempo em que o aluno foi exposto a um ambiente educacional (FERRÃO, 2003; RAUDENBUSH, 2004b). Esse indicador pressupõe uma avaliação inicial e final, a partir da qual se pode estimar a mudança ocorrida.

Embora o Enade não adote um levantamento longitudinal avaliando o mesmo estudante quando ingressa e conclui um curso, ele avaliou, em um conjunto de edições, estudantes ingressantes e concluintes no mesmo ano. Assim, produziu dados mais próximos do que era possível com o Exame

Nacional de Cursos (ENC), para se tentar chegar a uma medida aproximada desse valor agregado. Assim, se for razoável assumir que o nível de desempenho dos alunos concluintes era, na época em que ingressaram, similar ao dos alunos ingressantes no ano em que se coletou os dados, então as diferenças médias ingressantes-concluintes, ainda que obtidas no mesmo ano, poderiam ser tomadas como indicativos de mudança típica (média). Essa é a lógica que fundamenta o Índice de Diferença de Desempenho (IDD). Com esse índice, calcula-se a média esperada do desempenho dos concluintes do curso considerando alguns preditores baseados no perfil do ingressante: desempenho médio no Enade ou no Exame Nacional do Ensino Médio (Enem) e proporção de estudantes ingressantes cujos pais tenham nível de escolaridade superior. A diferença entre o desempenho esperado, a partir do perfil prévio dos alunos que ingressam no curso, e o valor, de fato, observado entre os concluintes é o IDD.

O controle estatístico do IDD foi um avanço importante em relação ao ENC, pois tornou os índices mais consistentes e menos enviesados por variáveis irrelevantes à avaliação do curso (diferenças iniciais entre os estudantes). O IDD consiste em uma tentativa de solucionar o primeiro problema de validade, visto que, por meio do controle estatístico, isola-se a variância confiável do construto relevante de outros fatores irrelevantes à avaliação do curso (MESSICK, 1995).

Distinção dos cinco níveis diferenciados de qualidade

A segunda questão, sobre se há um montante suficiente para se criar uma distinção em cinco níveis, ainda não foi testada. Nos documentos técnicos acerca da proposição desses níveis, não há dados estatísticos que isolem os fatores relevantes e irrelevantes reportando a quantidade de variância explicada por cada um deles.

Neste trabalho foi adotada uma abordagem de análise de regressão hierárquica multinível, de forma a acomodar uma estrutura dos dois níveis já mencionados (estudantes e cursos), buscando-se identificar qual parcela de variância está associada a fatores irrelevantes e, conseqüentemente, qual parcela efetiva poderia ser usada na criação de um índice de qualidade dos cursos superiores. Com isso, pretende-se verificar se essa medida é suficiente para criar uma distinção de cinco níveis, como o atual sistema propõe.

Resumidamente, em um sistema válido de avaliação de cursos, espera-se que haja uma parte suficiente de variância associada às diferenças que são construídas durante o percurso acadêmico do aluno na IES. Esse será o propósito principal deste estudo: identificar qual percentual de variância da prova do Enade seria potencialmente útil na avaliação dos cursos, pois este está associado ao que se pretende medir. Portanto, este estudo pretende identificar as potencialidades e as limitações da prova do Enade em detectar competências, habilidades e conhecimentos que são aprendidos no contexto específico do ensino superior e, por conseguinte, a capacidade destes de se tornarem um indicador da qualidade dos cursos.

Método

Participantes

Com o objetivo de testar o modelo de análise, selecionou-se o banco de dados que contém os resultados dos estudantes do curso de Administração no Enade de 2006, além das informações do questionário socioeconômico. Esses dados representam uma amostra de 130.266 estudantes de 1.121 cursos de Administração no país, 71.838 ingressantes e 58.428 concluintes, o que corresponde aproximadamente à totalidade de cursos oferecidos na época da avaliação. Do total de participantes, 52% são do sexo feminino e têm idade média de 26 anos. O desvio-padrão foi de sete anos. Para este estudo, foram selecionados os estudantes que responderam pelo menos 70% das questões da parte específica da prova.

Em relação à amostra de IES, aproximadamente a metade está localizada no Sudeste (49%), seguida do Sul (22%) e do Nordeste (14%). A grande maioria das IES analisadas fazem parte da rede particular de ensino (90,2%).

Materiais

Foi utilizada a prova do Enade de 2006, a qual apresenta 40 questões que avaliam conhecimentos, competências profissionais e habilidades acadêmicas de estudantes do ensino superior. Dessas questões, 10 (8 objetivas e 2 dissertativas) são relativas ao componente de formação geral (FG) e 30 (24 objetivas e 6 dissertativas) são específicas da área do curso em avaliação – componente específico (CE).

Foi analisado, também, o questionário socioeconômico respondido pelos estudantes, composto por 114 questões sobre: perfil socioeconômico, dedicação às atividades acadêmicas, hábitos de leitura, conhecimento de línguas, história acadêmica, uso do computador, acesso à internet e percepção de elementos da IES e da formação por ela oferecida. Por intermédio desse questionário, é possível extrair informações importantes a respeito do perfil dos estudantes, além de obter dados contextuais característicos dos diferentes cursos. O questionário foi preenchido pelos alunos em casa e entregue por eles no dia da realização da prova. Verificou-se que 80% dos estudantes responderam a ele.²

Especificamente em relação às variáveis socioeconômicas do questionário do Enade, partiu-se da análise realizada por Silva et al. (2010), com base no Enade de 2005, para congregar grupos de itens que se esperavam estar associados ao desempenho do estudante no exame. Esses autores se basearam na análise fatorial do questionário, nas correlações dos itens com o desempenho no componente específico da prova e na análise qualitativa das questões a fim de agrupá-las e assim utilizá-las de forma mais sintetizada em estudos sobre a influência das instituições na qualidade da educação.

O quadro 1 descreve as diferentes variáveis inseridas nos modelos estatísticos construídos. A variável “escolaridade dos pais” foi inserida durante o processo de modelagem, mas descartada em etapas posteriores por não mais apresentar um efeito significativo.

2 Para uma análise detalhada do questionário socioeconômico, ver Silva et al. (2010).

Quadro 1 Escalas das variáveis explicativas do modelo de regressão multinível

NÍVEL	VARIÁVEL	DESCRIÇÃO	CODIFICAÇÃO
1	Tipo_instit	Tipo de escola	1 – Pública 2 – Mista 3 – Privada
	escol_pais	Escolaridade dos pais	Variação de 1 a 5 (nenhuma escolaridade – ensino superior)
	renda	Renda familiar	Variação de 1 a 7 (até 3 salários mínimos – mais de 30 salários mínimos)
	Sus_trab	Formas de sustento e trabalho	Variação de 1 a 5 (não trabalha – principal responsável pelo sustento)
2	curso	Variável randômica associada ao código de cada curso	1.121 valores correspondentes aos cursos
	ce.measure_mean.0	Pontuação no CE ³	- 3,0 a 3,0
	nt_fg_mean.0	Pontuação na FG ⁴	0 a 100

Fonte: autoria própria.

³ Desempenho no CE na escala theta de habilidade (ingressantes).

⁴ Desempenho na FG (ingressantes).

Procedimento e modelagem estatística

Toda variância sistemática associada ao período em que os estudantes estavam sob influência do curso pode potencialmente se caracterizar como variância confiável de construto relevante (MESSICK, 1995). Após a separação desse último componente e, portanto, depois de expurgados os fatores irrelevantes, buscou-se investigar a quantidade de variância restante e que potencialmente poderia ser usada para avaliar os cursos e refazer as médias corrigidas, comparando-se os resultados dos cursos com e sem o controle de variáveis irrelevantes.

Na modelagem estatística foi usada a Análise de Regressão Hierárquica ou Modelos Multiníveis (RAUDENBUSH; BRYK, 2002; SNIJDERS; BOSKER, 2003; FERRÃO, 2003), empregando-se o programa de computador MLwiN.⁵ Adotou-se um modelo de dois níveis: os estudantes no nível 1 e os diferentes cursos no nível 2.

Empregou-se como variável dependente o resultado do Enade dos estudantes que concluíram a parte específica da prova. Como fatores explicativos do perfil prévio dos estudantes (nível 1), selecionaram-se as variáveis indicadoras de níveis socioeconômicos (se os estudantes trabalhavam para conseguir se sustentar, escolaridade dos pais e renda familiar). Essas variáveis foram incluídas no modelo como fatores fixos centralizados na média geral, que expurgam parte da variância associada a diferenças prévias de perfil socioeconômico dos estudantes; certamente importantes, mas irrelevantes para a avaliação da qualidade dos cursos. No nível 2, a variável “curso” configurou-se na especificação de agrupamentos de alunos de cada curso, configurando-se em um termo constante e randômico, correspondente à expectativa de desempenho médio dos alunos de cada curso quando todas as variáveis de controle do nível 1 fossem iguais a zero. Assim, no termo constante, há a representação da variabilidade do desempenho médio dos estudantes de cada curso expurgando-se as diferenças prévias decorrentes dos fatores socioeconômicos.

⁵ Disponível em: <<http://www.bristol.ac.uk/cmm/software/mlwin/>>.

Ainda no nível 2, buscou-se explicar as variabilidades dessas médias com base nos resultados médios no Enade dos ingressantes de cada IES. Assim, subtrai-se da variabilidade das médias dos cursos aquelas diferenças do desempenho médio dos ingressantes. Os resíduos desse termo constante no nível 2 consistem na parcela de variância que poderia ser atribuída às eventuais diferenças de qualidade entre cursos. Com isso, buscou-se verificar a proporção de variância relevante disponível no desempenho dos estudantes. Assim, é possível decompor a variância do desempenho do estudante concluinte no Enade separando-se a parcela mais relevante na construção de indicadores de qualidade dos cursos. A última etapa dessa análise consistiu em estimar os resíduos no nível 2 do modelo, que foram constituídos pelas notas mais refinadas dos cursos, eliminando-se parcelas irrelevantes da variância. Tendo essas notas, foi feito um estudo correlacional com os conceitos dos cursos calculados pelo método tradicional, a fim de se discutir os aspectos referentes à validade do uso do Enade para verificar a qualidade dos cursos.

Resultados

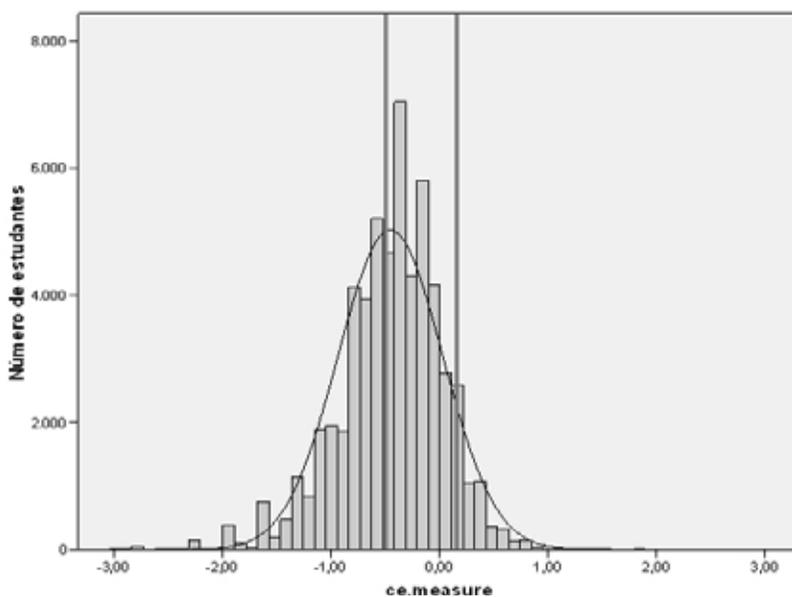
Para a análise dos dados, foi considerado como variável dependente o desempenho dos alunos concluintes da parte relativa ao CE da prova. Primeiramente foram eliminados da análise os estudantes que responderam menos de 23 questões, a fim de selecionar apenas aqueles que responderam 75% dos itens de CE. Utilizou-se o resultado do Enade dos estudantes concluintes do CE estimado com a Teoria de Resposta ao Item (TRI), cujos detalhes são apresentados em Silva et al. (2011). Essa escala é similar à escala Z, com valores típicos de -4 a +4, com média 0. Entretanto, em razão do procedimento de calibração empregado, a média 0 foi centrada na escala de dificuldades dos itens. Portanto, o valor 0 corresponde à parte dos itens com dificuldade média na prova. A escala de habilidade dos estudantes é relativa a esse ponto, o que faz com que a média não seja 0 porque a média de acertos é menor do que 50%. Na tabela 1 e no gráfico 1, são apresentadas as estatísticas descritivas e a distribuição das notas dos alunos na prova de CE.

Tabela 1 Estatísticas descritivas dos estudantes na prova de CE por momento do curso

MOMENTO/ESTATÍSTICAS	
INGRESSANTES	
Número total	71.496
Média	-0,74
Desvio-padrão	0,54
Mínimo	-4,87
Máximo	3,02
CONCLUINTES	
Número total	58.066
Média	-0,46
Desvio-padrão	0,48
Mínimo	-4,84
Máximo	2,26

Fonte: autoria própria.

Gráfico 1 Distribuição de habilidades no componente específico (concluintes)



Fonte: autoria própria

As variáveis independentes foram centralizadas a fim de permitir a comparação entre elas. Primeiramente, foi analisado o modelo nulo, sem variáveis explicativas, para identificar qual parte da variância total poderia ser atribuída aos níveis 1 e 2. A notação formal do modelo é a seguinte:

$$\begin{aligned}y_{ij} &= \beta_{0j} + e_{ij} \\ \beta_{0j} &= \beta_0 + \mu_{0j} \\ \mu_{0j} &\sim N(0, \sigma_{\mu j}^2) \\ e_{ij} &\sim N(0, \sigma_e^2)\end{aligned}$$

A notação y_{ij} representa o desempenho na parte específica do Enade do estudante i no curso j . O desempenho é previsto por um parâmetro β_{0j} , composto pela média geral da amostra β_0 e pela média dos concluintes do curso j , do qual o estudante i é parte; μ_{0j} representa o desvio da IES em relação à média geral e mais um resíduo e_{ij} , relacionado ao estudante, que indica o quanto ele se afasta da previsão com base na média da IES.

Assim, esse modelo separa a variância total em dois componentes: *i*) a variância dos estudantes com relação à média de seu curso σ_e^2 , que pertence ao nível 1; e *ii*) a variância da média dos estudantes de uma IES j com relação à média geral $\sigma_{\mu j}^2$, que pertence ao nível 2.

Vale salientar que o modelo tenta explicar a variância total ($\sigma_{\mu j}^2 + \sigma_e^2$) empregando os parâmetros da IES no nível 2. Assim, serve para verificar a variância explicada $\sigma_{\mu j}^2$ versus a não explicada σ_e^2 , o que permite calcular o coeficiente de correlação intraclasse (r_{ICC}) ao dividir-se a variância do nível macro (IES) pela variância total. Portanto, ela pode variar de 0 a 1, com base na fórmula:

$$r_{ICC} = \frac{\sigma_{\mu j}^2}{\sigma_{\mu j}^2 + \sigma_e^2}$$

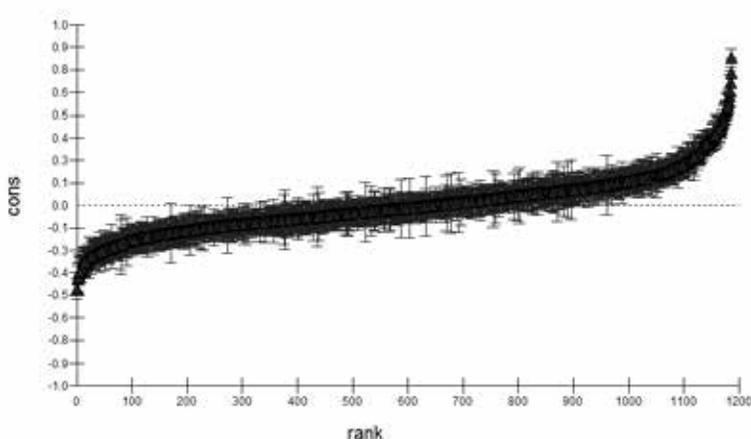
O valor nulo ou próximo de zero indica que as instituições são homogêneas entre si e que o desempenho escolar do aluno é independente do local que ele frequenta (FERRÃO; FERNANDES, 2003; JESUS; LAROS, 2004). Indica ainda que a análise multinível não precisa ser aplicada. No caso do presente estudo, tal resultado mostraria que a diferença nos desempenhos dos

estudantes não pode ser atribuída necessariamente à IES frequentada pelo estudante.

Portanto, nesse modelo, no nível 1 ($N = 45.550$), estão os estudantes agrupados nos cursos das IES brasileiras que participaram do Enade 2006, concluintes do curso de Administração, que, por sua vez, constitui o nível 2 (1.121 cursos). De acordo com o modelo nulo, a estimativa da variância entre as instituições é de 0,031. A variância entre os estudantes dentro das instituições é igual a 0,204 e a variância total é de 0,235. Desse modo, foi possível calcular o coeficiente de correlação intraclasse, que mede o quanto a variação total decorre das diferenças entre as instituições. Segundo o resultado observado, somente 13,2% da variação total se deve à variação entre os cursos dessas instituições.

O gráfico 2 mostra os valores do parâmetro das 1.121 instituições, ordenados do menor para o maior, de modo que é possível visualizar a existência de pouca variabilidade entre os cursos. Uma parcela maior da variância ocorre dentro dos cursos. Pode-se inferir que há efeito pequeno dos *clusters* formados pelas IES no desempenho dos estudantes.

Gráfico 2 Ranking das instituições no componente específico, no modelo nulo



Fonte: autoria própria.

Na etapa seguinte, o modelo foi ampliado incluindo-se as variáveis do nível 1 (nível socioeconômico do aluno, por exemplo) e do nível 2 (média

no Enade dos estudantes ingressantes nos cursos), com o objetivo de remover a variância irrelevante para a avaliação da instituição. A notação geral do modelo é:

$$\begin{aligned} y_{ij} &= \beta_{0j} + (X_{1ij} + \dots + X_{vij}) + e_{ij} \\ \beta_{0j} &= \beta_0 + (\beta_{1j} + \dots + \beta_{vj}) + \mu_{0j} \\ \mu_{0j} &\sim N(0, \sigma_{\mu_j}^2) \\ e_{ij} &\sim N(0, \sigma_e^2) \end{aligned}$$

Nesse modelo, os termos $X_{1ij} + \dots + X_{vij}$ indicam as variáveis X_{vij} predictoras consideradas no nível 1. Já $\beta_{1j} + \dots + \beta_{vj}$ indicam as variáveis predictoras β_{vj} do nível 2. É importante ressaltar que a variância total dos termos nesse modelo ($\sigma_{\mu_j}^2 + \sigma_e^2$) é menor que no modelo nulo, e a magnitude da redução indica o montante de variância explicado pelas variáveis que foram inseridas:

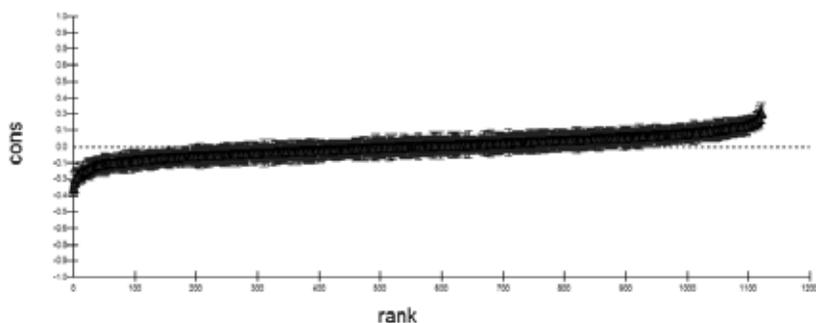
$$r = \frac{\sigma_{Tot_nulo}^2 - \sigma_{Tot_odt}^2}{\sigma_{Tot_nulo}^2}$$

Como já dito, esse modelo inclui, *a priori*, as principais variáveis do nível 2 relacionadas ao perfil cognitivo dos estudantes, a saber: o desempenho no CE e na FG como medidas do perfil típico médio dos ingressantes na IES. Assim, é feito o controle do desempenho inicial no nível 2.

Nesse modelo, a variância residual das médias das IES ($\sigma_{\mu_j}^2$) representa uma parcela de variância mais 'pura', pois remove parte da variância que não poderia ser atribuída à qualidade da IES, que compreende fatores externos relacionados ao perfil de sua clientela. Verificou-se que a variância residual das médias das IES reduziu-se de 13,2% (do modelo nulo) a 5% após inclusão dos desempenhos médios dos ingressantes no modelo. Em seguida, com a inclusão das variáveis predictoras do nível 1 (escolaridade dos pais, renda familiar, formas de sustento e trabalho), a variância foi reduzida a 4%.

Ao visualizar o gráfico 3, semelhante ao 2 (refeito com base no último modelo), é possível verificar que a variabilidade entre os cursos é ainda menor no modelo final. Desse modo, é possível inferir que as diferenças se dão mais entre os alunos do que entre os cursos.

Gráfico 3 Ranking das instituições no componente específico após controle



Fonte: autoria própria.

Assim é possível analisar os efeitos das variáveis explicativas de ambos os níveis. Os resultados encontrados podem ser vistos na tabela 2, que apresenta as estatísticas referentes ao modelo nulo e final encontrado neste estudo.

Tabela 2 Modelo sem variáveis explicativas criado após a inserção das variáveis de controle

VARIÁVEIS EXPLICATIVAS	MODELO NULO (SEM VARIÁVEIS EXPLICATIVAS)			MODELO DE REFERÊNCIA (APÓS INSERÇÃO DE VARIÁVEIS DE CONTROLE)		
	EFEITO	EP	RAZÃO TOTAL	EFEITO	EP	RAZÃO TOTAL
Efeitos Fixos						
Intercepto (β_0)	-0,467			-0,453		
ce.measure_mean.0 (β_{1j})				0,504	0,035	14,4
nt_fg_mean.0 (β_{2j})				0,009	0,001	9
Renda (X_{1ij})				0,029	0,002	14,5
Sus_trab (X_{2ij})				0,055	0,002	27,5

VARIÁVEIS EXPLICATIVAS	MODELO NULO (SEM VARIÁVEIS EXPLICATIVAS)			MODELO DE REFERÊNCIA (APÓS INSERÇÃO DE VARIÁVEIS DE CONTROLE)		
	EFEITO	EP	RAZÃO TOTAL	EFEITO	EP	RAZÃO TOTAL
Nível 2 (variância de μ_{0j} $\sigma_{\mu j}^2$)	0,031			0,011		
Nível 1 (variância de e_{ij} σ_e^2)	0,204			0,201		
Correlação intraclasses		0,13			0,04	
Deviance		72969,06			57251,29	
Número de parâmetros estimados		3			7	
Diferença nos deviances					3929,444	
Diferença do número de parâmetros					4	

Fonte: autoria própria.

Uma última etapa consistiu em analisar as diferenças de desempenho dos estudantes segundo os conceitos do Enade, antes e após o controle das variáveis pelo modelo final da análise multinível (tabela 3). Antes do controle, há, como esperado, uma diferença marcante entre os cursos das IES associados ao seu conceito, uma vez que cursos com conceito mais alto apresentaram médias mais altas. No entanto, depois do controle, as diferenças entre os cursos com conceitos distintos diminuem radicalmente e se tornam muito pequenas. Assim, grande parte das diferenças entre os cursos pode ser atribuída ao perfil socioeconômico e cognitivo dos alunos que ingressam em cada IES.

Tabela 3 Diferenças de desempenho dos estudantes por conceito antes e após controle de variáveis externas

CONCEITO	ANTES DO CONTROLE		APÓS O CONTROLE	
	MÉDIA	ERRO-PADRÃO DA MÉDIA	MÉDIA	ERRO-PADRÃO DA MÉDIA
1	-0,84	0,017	-0,05	0,019
2	-0,63	0,004	-0,02	0,004
3	-0,45	0,003	0,00	0,002
4	-0,24	0,006	0,02	0,006
5	0,02	0,009	0,04	0,010

Fonte: autoria própria.

Discussão

Este estudo decompôs a variância do desempenho no Enade do estudante concluinte separando a parcela mais relevante para a construção de indicadores de qualidade dos cursos. O resultado mais importante indica que a maior parte da variabilidade das notas do Enade está dentro dos cursos, e não entre os cursos. Verificou-se que a correlação intraclasse diminuiu de 0,13 para 0,04 depois da inserção das variáveis de controle. Esse valor é compatível ou até menor que o valor da correlação intraclasse de outros estudos similares (FERRÃO; FERNANDES, 2003; JESUS; LAROS, 2004; SOARES, 2004; SOARES et al., 2001, 2004), apesar da ressalva de que foram feitos no âmbito do ensino fundamental e não do ensino superior.

Portanto, os resultados sugerem que, antes do controle, grande parte da variabilidade das médias entre as instituições que irão compor as atuais notas dos cursos (de 1 a 5) é explicada pelos perfis socioeconômicos e o desempenho prévio dos estudantes. Assim, o indicador de qualidade com base na média dos concluintes é em grande parte confundido com as condições prévias dos estudantes, o que dificulta identificar a parcela que pode ser atribuída à influência do curso (SOARES, RIBEIRO; CASTRO, 2001; PRIMI, 2006a, 2006b). Tais dados apontam a relevância de se controlar estatisticamente o efeito dessas variáveis, caso contrário se atribuiria aos cursos e, assim, também às instituições um efeito que não cabe a estes (FERRÃO, 2003; FERRÃO; FERNANDES, 2003, PRIMI, 2006a, 2006b; BROOKE; SOARES, 2008).

A variabilidade das médias dos cursos, decorrentes dos fatores socioeconômicos e do desempenho prévio, que foi o foco de análise deste artigo, é um componente equivalente ao IDD. As variáveis de controle usadas e o método (multinível *versus* regressão com dados agregados) não são idênticos, mas são índices bastante similares. Uma questão importante neste estudo, e que também se refere ao IDD, é se 4% da variabilidade total do desempenho do Enade é um montante suficientemente útil para se criar um índice que diferencia a qualidade dos cursos. Um ponto importante nessa comparação é que, no cálculo da correlação intraclasse que acarreta essa proporção, o denominador refere-se à variância no Enade entre indivíduos, e o numerador à variância entre médias de cursos. Este apresenta resultado naturalmente menor do que aquele. Ao considerar-se que as amostras

de números de alunos de cada curso vêm de uma mesma população, a variabilidade das médias seria dada por $\sigma\sqrt{N}$. Essa diferença de escalas entre a média dos cursos (nível 2) e a média dos estudantes (nível 1) geralmente não é considerada na discussão das correlações intraclasse, mas é certamente importante na análise psicométrica desse índice, ao ser julgado se esse montante é útil ou não.

Desse modo, a nota final do Enade, com base no desempenho dos concluintes sem controle, contém uma parcela considerável de variância irrelevante que pode comprometer a validade do índice. Este estudo ressalta a importância de índices, como o IDD e, alternativamente, destaca a abordagem multinível como forma de construir índices potencialmente mais úteis para a avaliação da qualidade dos cursos. Seria importante para estudos futuros e para a avaliação da validade do IDD a análise da correlação desse índice com outros indicadores mais diretos da qualidade dos cursos, bem como análises econométricas de identificação do impacto dos cursos e do quanto esse se relaciona ao IDD.

Referências

AERA; APA; NCME. *Standards of Educational and Psychological testing*. Washington, DC: American Educational Research Association, 1999.

ANTONAKIS, J. et al. On making causal claims: a review and recommendations. *The Leadership Quarterly*, v. 21, n. 6, p. 1086-1120, 2010.

BROOKE, N.; SOARES, J. F. (Org.). *Pesquisa em eficácia escolar: origem e trajetórias*. Belo Horizonte: Editora UFMG, 2008.

CARROLL, J. B. A model of school learning. *Teachers College Record*, v. 64, p. 723-733, 1963.

FERRAO, M. E. *Introdução aos modelos de regressão multinível em educação*. Campinas, São Paulo: Komedi, 2003.

FERRAO, M. E.; FERNANDES, C. O efeito-escola e a mudança – dá para mudar? Evidências da investigação brasileira. *Revista Iberoamericana sobre Calidad, Eficacia y*

Cambio en Educaci3n (Reice), v. 1, n. 1, 2003. Dispon3vel em: <<http://www.ice.deusto.es/rinace/reice/vol1n1/FerrayFernandes.pdf>>. Acesso em: 8 jun. 2009.

GOLDSTEIN, H. The difficulty of ranking schools: the limits to 'value-added'. *New Economy*, v. 27, n. 4, p. 391-405, 2001.

GOLDSTEIN, H. et al. A multilevel analysis of school examination results. *Oxford Review of Educacion*, v. 19, n. 4, p. 425-433, 1993.

GOLDSTEIN, H.; WOODHOUSE G. Pesquisa sobre efic3cia escolar e politicas educacionais. In: BROOKE, N.; SOARES, J. F. (Org.). *Pesquisa em efic3cia escolar: origem e trajet3rias*. Belo Horizonte: Editora UFMG, 2008. p. 411-424.

JESUS, G. R.; LAROS, J. A. Efic3cia escolar: regress3o multin3vel com dados de avalia3o em larga escala. *Avalia3o Psicol3gica*, v. 3, n. 2, p. 93-106, 2004.

LANDEIRA-FERNANDEZ, J.; PRIMI, R. Compara3o do desempenho entre calouros e formandos no Prov3o de Psicologia 2000. *Psicologia: Reflex3o e Cr3tica*, v. 15, n. 1, p. 219-234, 2002.

MESSICK, S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, v. 50, n. 9, p. 741-749, 1995.

PRIMI, R. Evid3ncias de validade das provas do Enade – 2004. In: RISTOFF, D.; LIMANA, A.; BRITO, M. R. (Org.). *Enade: perspectiva de avalia3o din3mica e analise de mudan3as*. Bras3lia: Instituto Nacional de Estudos e Pesquisas Educacionais An3sio Teixeira (Inep), p. 59-76, 2006a.

PRIMI, R. A validade do *Enade para avalia3o da qualidade dos cursos de institui3es de ensino superior*. Projeto de Pesquisa. Itatiba: Universidade S3o Francisco, LabAPE, 2006b.

PRIMI, R.; FERNANDEZ, J.L.; ZIVIANI, C. O prov3o de psicologia: objetivos, problemas, consequ3ncias e sugest3es. *Psicologia: Teoria e Pesquisa*, v. 19, n. 2, p. 109-116, 2003.

PRIMI, R.; MUNIZ, M.; NUNES, C. H. S. S. Defini3es contempor3neas de validade de testes psicol3gicos. In: Claudio Simon Hutz. (Org.). *Avan3os e pol3micas em avalia3o psicol3gica*. S3o Paulo: Casa do Psic3logo, p. 243-265, 2009.

RAUDENBUSH, S. W. *Schooling, statistics, and poverty: can we measure school improvement?* New Jersey: Educational Testing Service, 2004a.

_____. What are value-added models of estimating and what does this imply for statistical practice. *Journal of Educational and Behavioral Statistics*, v. 29, n. 1, p. 121-129, 2004b.

RAUDENBUSH, S. W.; BRYK, A. S. *Hierarchical linear models: applications and data analysis methods (advanced quantitative techniques in the social sciences)*. London: Sage Publications, 2002.

RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, v. 66, n. 5, p. 688-701, 1974.

SILVA, M. C. R. et al. Teoria de Resposta ao Item: equalização da prova de Administração do Enade. In: CAPOVILLA, F. C. (Org.). *Transtornos de aprendizagem: progressos em avaliação e intervenção preventiva e remediativa*, v. 1. São Paulo: Memnon, 2010. p. 239-246.

SNIJEDERS, T. A. B.; BOSKER, R. *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London: Sage Publications, 2003.

SOARES, J. F. O efeito da escola no desempenho cognitivo de seus estudantes. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educacion (Reice)*, v. 2, n. 2, p. 83-104. 2004. Disponível em: <<http://www.ice.deusto.es/rinace/reice/vol2n2/Soares.pdf>>. Acesso em: 20 mai. 2008.

_____. O efeito da escola no desempenho cognitivo de seus alunos. In: SOUZA, A. M. (Org.). *Dimensões da avaliação educacional*. Petrópolis: Vozes, 2005, p. 174-204.

SOARES, J. F., CASTRO, C. M.; RIBEIRO, L. M. *O Provão: os cursos A são os que mais oferecem aos estudantes?* Belo Horizonte: UFMG. Texto não publicado, 2000.

SOARES, J. F., RIBEIRO, L.; CASTRO, C. M. Valor agregado de instituições de ensino superior em Minas Gerais para os cursos de Direito, Administração e Engenharia Civil. *Dados*, v. 44, n. 2, p. 363-396, 2004.

URBINA, S. *Essentials of psychological testing*. New Jersey: John Wiley & Sons Inc., 2004.

Ricardo Primi

Doutor em Psicologia Escolar e do Desenvolvimento Humano pela Universidade de São Paulo.

Professor da Universidade São Francisco, Itatiba (SP), Brasil;

ricardo.primi@saofrancisco.edu.br

Marjorie Cristina Rocha da Silva

Doutora em Psicologia na Universidade de São Francisco.

Professora do Centro Universitário FIEO, UNIFIEO, Osasco (SP), Brasil;

silvamarjorie@yahoo.com.br

Daniel Bartholomeu

Mestre e Doutor na área de Avaliação Psicológica na Universidade de São Francisco.

Professor do Centro Universitário Salesiano de São Paulo (Unisal), Americana (SP), Brasil;

d_bartholomeu@yahoo.com.br

Artigo submetido em 07/06/2017

Aprovado em 26/04/2018

UM ESTUDO DA RELAÇÃO ENTRE TRAÇO LATENTE E VARIÁVEIS CONTEXTUAIS NO SAEB E ENEM

A STUDY RELATING LATENT TRAITS TO CONTEXTUAL VARIABLES FOR THE
SAEB AND ENEM EXAMS

UN ESTUDIO DE LA RELACIÓN ENTRE TRAZO LATENTE Y VARIABLES
CONTEXTUALES EN LOS EXAMENES SAEB Y ENEM

Flávio Bambirra Gonçalves

Bárbara da Costa Campos Dias

RESUMO

Este artigo tem como objetivo aplicar uma metodologia proposta recentemente para estudar a relação entre o traço latente (proficiência) e as variáveis presentes em questionários contextuais em avaliações educacionais de larga escala no Brasil. A metodologia considerada utiliza a Teoria de Resposta ao Item para modelar as respostas dadas aos testes e modela os traços latentes hierarquicamente por meio de um modelo de regressão linear com covariáveis advindas dos questionários. Os dados faltantes nos questionários contextuais são tratados estatisticamente através de sua inclusão no modelo considerado. Uma abordagem bayesiana é considerada, o que permite, além de estimar as proficiências e os parâmetros dos itens, a estimação das respostas faltantes. Além disto, a abordagem conjunta considerada permite a quantificação de toda a incerteza de maneira eficiente. São apresentadas análises utilizando-se os dados do Sistema de Avaliação da Educação Básica (Saeb) e o Exame Nacional do Ensino Médio (Enem). Os resultados evidenciam tanto a relação entre traço latente e as variáveis contextuais quanto a existência de padrões de não resposta nos questionários.

Palavras-chave: Teoria de Resposta ao Item; inferência Bayesiana; modelos de regressão; dados faltantes.

ABSTRACT

This paper considers a recently proposed methodology to analyse the relation between latent traits and covariates from contextual questionnaires in large scale educational assessment exams in Brazil. The methodology uses Item Response theory to model the answers given in a test and hierarchically models the latent traits in a regression model with covariates from the questionnaires. Missing data from those questionnaires are included in the statistical model considered and the Bayesian approach adopted allows the estimation of the missing data. Furthermore, the simultaneous approach of the methodology efficiently quantifies all the uncertainty involved in the analysis. Data sets from the main two large scale exams in Brazil are considered – Saeb and Enem. Results indicate considerable relation between latent traits and covariates and not at random missing data patterns.

Keywords: Item Response Theory; Bayesian inference; regression models; missing data.

RESUMEN

Este artículo considera una metodología recientemente propuesta para analizar la relación entre los trazos latentes y las covariables de los cuestionarios contextuales en los exámenes de evaluación educacional a gran escala en Brasil. La metodología utiliza Item Response Theory para modelar las respuestas dadas en una prueba y modela jerárquicamente los trazos latentes en un modelo de regresión con covariables de los cuestionarios. Los datos faltantes de esos cuestionarios se incluyen en el modelo estadístico considerado y el enfoque Bayesiano adoptado permite, entre otras cosas, estimar los datos faltantes. Además, el enfoque simultáneo de la metodología cuantifica de manera eficiente toda la incertidumbre involucrada en el análisis. Se consideran conjuntos de datos de los dos principales exámenes a gran escala en Brasil: Saeb y Enem. Se encuentran resultados interesantes considerando la relación entre los rasgos latentes con las covariables y los patrones de datos faltantes.

Palavras clave: Teoría de Respuesta al Item; inferencia Bayesiana; modelos de regresión; datos faltantes.

Introdução

As avaliações educacionais são um dos principais instrumentos para se coletar informações sobre o desempenho e os resultados do sistema educacional, como, por exemplo, os fatores que afetam o processo de aprendizagem, a qualidade das escolas e os outros diversos fatores que podem descrever e auxiliar no desenvolvimento de políticas que visam a melhoria do processo de aprendizagem.

O Sistema de Avaliação da Educação Básica (Saeb) e o Exame Nacional do Ensino Médio (Enem), organizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), são as principais avaliações em larga escala do Brasil. O Saeb aplica testes que avaliam o desempenho escolar nas áreas de Língua Portuguesa e Matemática e esses testes são aplicados a alunos matriculados no 5º ano e 9º ano do ensino fundamental e na 3ª série do ensino médio. O Enem é uma prova elaborada para verificar o domínio de competências e habilidades dos estudantes que estão concluindo ou que já concluíram o ensino médio, e o seu resultado utilizado para o ingresso em diversas universidades no país.

No contexto de avaliações educacionais, os questionários contextuais são uma importante ferramenta para se analisar os fatores que podem influenciar no processo de aprendizagem dos alunos. Um dos principais objetivos é investigar possíveis fatores relacionados às proficiências acadêmicas, sendo o questionário uma rica fonte de informação sobre esses fatores, por exemplo, permitindo traçar um perfil socioeconômico e cultural dos estudantes. De acordo com Franco et al. (2003), os questionários contextuais desempenham o papel de oferecer fatores explicativos para a modelagem das proficiências dos alunos, medidas pelos instrumentos cognitivos. Segundo Laros e Marciano (2008), a identificação de variáveis que influenciam a proficiência em Língua Portuguesa é essencial para a elaboração de métodos que visam melhorar o desempenho acadêmico dos alunos.

Esses questionários são aplicados conjuntamente com a avaliação e coletam informações referentes aos aspectos da vida escolar, do nível socioeconômico, da classe social, cultural etc. Por exemplo, pesquisadores acreditam existir uma forte relação entre desempenho escolar e classe social, indicando que desigualdades nos resultados escolares

têm origem na família e no contexto social das escolas (COLEMAN, 1966; JENCKS, 1972).

Uma característica recorrente na aplicação dos questionários, porém, é a presença de dados faltantes, ou seja, nem todas as perguntas são respondidas por todos os alunos. Além disso, esses dados faltantes aparecem, tipicamente, sem padrão de não resposta. Esta questão é ainda mais relevante em um contexto, no qual se pretende utilizar as informações contidas nas respostas dos questionários para se identificar possíveis fatores que influenciem no rendimento dos alunos. Por isso, é indispensável um tratamento estatístico adequado dos dados faltantes para se obter resultados confiáveis na análise.

Recentemente, Gonçalves e Dias (2018) propuseram uma metodologia estatística eficiente para analisar dados provenientes de avaliações educacionais, utilizando as informações dos questionários contextuais para identificar possíveis fatores relacionados ao desempenho dos alunos. Além disso, a metodologia proposta considera a possibilidade de se ter respostas faltantes nos questionários de forma que toda a informação disponível é utilizada, o que possibilita a estimação das respostas não observadas.

O objetivo do presente artigo é analisar dados de avaliações em larga escala no Brasil utilizando-se a metodologia proposta por Gonçalves e Dias (2018). Desta forma, dados provenientes dos questionários contextuais são utilizados para se investigar possíveis fatores relacionados ao desempenho dos alunos, mais especificamente, são considerados conjuntos de dados referentes ao Saeb e ao Enem. No caso do Saeb, consideram-se as provas de Português e Matemática aplicadas: *i*) ao 5º ano do ensino fundamental em 2015 nos estados de Minas Gerais e Bahia; *ii*) à 3ª série do ensino médio em 2015 nas regiões Sudeste e Nordeste. Para o Enem, consideram-se, também, as provas de Português e de Matemática do ano de 2015 nas regiões Sudeste e Nordeste.

Método

A metodologia proposta por Gonçalves e Dias (2018) é descrita de forma concisa nesta seção. Maiores detalhes podem ser encontrados na referência original. Os autores propõem um modelo estatístico da Teoria de Resposta ao Item (TRI), que modela as respostas dadas a um teste em função das características dos itens e da proficiência (habilidade, traço latente) dos alunos e considera as respostas dadas pelos alunos ao questionário contextual para identificar possíveis fatores relacionados aos seus respectivos rendimentos. Além disso, possíveis respostas faltantes nos questionários são adequadamente modeladas de forma a não prejudicar a análise.

Modelo

Seja Y_{ij} a variável indicadora de resposta correta do indivíduo j ao item i para $i = 1, \dots, I$ e $j = 1, \dots, J$, a probabilidade de esta resposta estar correta é modelada pelo clássico modelo probito de três parâmetros (3PNO):

$$P(Y_{ij} = 1 \mid \theta_j, a_i, b_i, c_i) = c_i + c_i \Phi(a_i \theta_j - b_i),$$

(1) onde $\Phi(\cdot)$ é a função distribuição da normal padrão, θ_j é a proficiência do indivíduo j e os parâmetros de discriminação, dificuldade e acerto casual do item i são denotados por a_i , b_i e c_i , respectivamente. Quanto maior for a proficiência do aluno, maior é a probabilidade de ele acertar um dado item.

Com o objetivo de investigar possíveis fatores contextuais que ajudam a explicar as proficiências dos alunos, define-se um segundo nível no modelo, no qual as proficiências são (estatisticamente) explicadas por covariáveis presentes no questionário contextual:

$$\theta_j = X_j \beta + e_j,$$

onde $\beta = (\beta_0, \beta_1, \dots, \beta_Q)'$ são os coeficientes da regressão, X_j é a j -ésima linha da matriz regressora X , que contém covariáveis extraídas do questionário contextual, e $e_j \sim N(0, \sigma_e^2)$ são erros aleatórios independentes. Os problemas de identificabilidade do modelo são contornados assumindo-se $\beta_0 = 0$ e $\sigma_e^2 = 1$. Esses valores fixam a escala das proficiências auxiliando, assim, a interpretação dos valores estimados.

As estimativas a serem obtidas para os coeficientes β são de particular interesse, uma vez que indicam a existência e a magnitude da relação entre a respectiva covariável do questionário e a proficiência.

Um terceiro nível do modelo considera a distribuição das respostas de todos os alunos às questões consideradas do questionário contextual. Isso inclui as questões não respondidas, o que, por sua vez, permite que seja feita a estimação dessas respostas bem como de relações entre a propensão a não responder e a resposta que seria dada. Em particular, pode-se comparar as respostas observadas com as não observadas (portanto, estimadas).

Inferência

Sob o Paradigma bayesiano, todo o processo de inferência é realizado baseado na distribuição *a posteriori* de todos os parâmetros e de todas as variáveis desconhecidas no modelo. Dada a complexidade dessa distribuição, Gonçalves e Dias (2018) propõem um algoritmo MCMC que permite explorá-la através de estimativas de Monte Carlo, que permitem realizar estimação pontual e intervalar das quantidades de interesse. Os detalhes do algoritmo são omitidos aqui e podem ser encontrados na referência original. As distribuições *a priori* são escolhidas de acordo com a escala dos parâmetros e são pouco informativas.

Análises

São analisados quatro conjuntos de dados referentes ao Saeb, todos do ano de 2015. Mais especificamente: 5º ano EF, Matemática/Português, Minas Gerais e Bahia; 3ª série EM, Matemática/Português, Sudeste e Nordeste. Considerando o Enem, são analisados dois conjuntos de dados referentes às provas de Matemática e Português do ano de 2015, nas regiões Sudeste e Nordeste. Em todas as análises, as provas de Português são analisadas separadamente (sem informação dos questionários contextuais) e as estimativas obtidas para as proficiências em Português são utilizadas, conjuntamente com algumas covariáveis do questionário, para explicar a proficiência em Matemática na análise dos dados referentes a esta prova.

Os conjuntos de dados do Saeb são referentes a todos os alunos que fizeram o CADERNO 1 (blocos 1 e 2), totalizando 22 itens no 5º ano e 26 na 3ª série de cada uma das duas provas. Além disso, são considerados os alunos que responderam, no mínimo, três itens na prova e preencheram o questionário. No caso do Enem, é selecionada uma amostra de 10 mil alunos de cada uma das duas regiões consideradas, dentre aqueles que fizeram a prova amarela.

As seguintes perguntas do questionário contextual foram consideradas em cada uma das análises referentes ao Saeb:

Q1 - Como você se considera?

a) Preto(a) ou Pardo(a) b) Indígena ou Amarelo(a) c) Branco(a)

Q2 - Qual o seu sexo?

a) Masc. b) Fem.

Q3 - Sua mãe, ou a mulher responsável por você, sabe ler e escrever?

a) Sim b) Não

No caso do Enem, são consideradas as seguintes perguntas:

Q1 - Em que tipo de escola você frequentou o Ensino Médio?

- a) Somente em escola pública
- b) Parte em pública e parte em privada
- c) Somente em escola privada

Q2 - Você exerce ou já exerceu atividade remunerada?

a) Não b) Sim

Q3 - Na sua residência tem acesso à internet?

a) Não

b) Sim

Como todas as covariáveis são categóricas, elas são incluídas no modelo de regressão através de variáveis indicadoras. A codificação das variáveis com três categorias é feita de maneira que a primeira covariável indicadora se refira à resposta a (coeficiente β_{11}) e a segunda à resposta b (coeficiente β_{12}). No caso das covariáveis com duas categorias, a covariável indicadora se refere à resposta a (coeficiente β_2 para a segunda covariável e β_3 para a terceira). A proficiência em Português é uma covariável contínua com respectivo coeficiente β_4 . A análise dos resultados do ajuste do modelo de regressão através das estimativas dos seus respectivos coeficientes deve ser feita em concordância com a escala fixada (variância 1). Além disso, as proficiências em Português são estimadas na escala (0,1).

A análise de padrões de não resposta é feita através da comparação das proporções de cada resposta observada e das respectivas proporções estimadas dentre os dados faltantes. Assume-se a resposta mais provável *a posteriori* como sendo a resposta que seria dada, caso a respectiva pergunta tivesse sido respondida. Nos casos em que existe informação disponível para estimação dos dados faltantes, diferenças significativas entre as proporções observadas e estimadas indicam a existência de padrões não aleatórios de não resposta, ou seja, há uma relação entre a resposta e a probabilidade daquela pergunta não ser respondida.

As duas fontes de informação para estimação dos dados faltantes são: *i*) respostas dadas a outras perguntas pelo mesmo aluno, no caso em que uma estrutura de dependência entre essas respostas é assumida no modelo; *ii*) a proficiência do aluno, estimada a partir das respostas do teste, desde que a covariável referente à resposta faltante tenha uma relação significativa (medida no modelo de regressão) com a proficiência, no caso em que se assume independência entre as respostas, como é assumido nas análises deste artigo; e *iii*) passa a ser a única fonte de informação, condicionada à existência da supracitada relação.

Resultados

Saeb, 5º ano – Minas Gerais e Bahia

Os conjuntos de dados analisados são formados por 2.021 alunos de Minas Gerais e 6.668 da Bahia. Entre os alunos de Minas Gerais, 329 não responderam a uma das três perguntas do questionário, 28 não responderam a duas e 34 não responderam a três, totalizando 487 dados faltantes. Para o estado da Bahia, estes números são 1.523, 132 e 73, totalizando 2.006 dados faltantes.

O ajuste dos modelos de regressão indica (ver tabela 1), em ambos os estados, uma forte associação positiva entre as proficiências em Português e em Matemática, além de um rendimento médio em matemática um pouco maior de alunos do sexo masculino. Além disso, não ficou evidenciada uma relação entre a proficiência em Matemática e a cor da pele, em ambos os estados. Para o estado de Minas Gerais, alunos cuja mãe ou mulher responsável sabe ler ou escrever tiveram uma proficiência média um pouco maior.

Tabela 1 Estimativa dos coeficientes da regressão - média *a posteriori* e intervalo de credibilidade de 99%, referentes aos conjuntos de dados do 5º ano EF do Saeb - Bahia (BA) e Minas Gerais (MG)

COEFICIENTES		ESTIMATIVA
BA	β_{11}	0.102 (-0.027, 0.227)
	β_{12}	0.078 (-0.124, 0.284)
	β_2	0.249 (0.154, 0.351)
	β_3	0.115 (-0.055, 0.295)
	β_4	1.265 (1.189, 1.344)
MG	β_{11}	-0.086 (-0.264, 0.092)
	β_{12}	-0.177 (-0.523, 0.163)
	β_2	0.317 (0.163, 0.474)
	β_3	0.496 (-0.047, 0.949)
	β_4	1.424 (1.296, 1.559)

Os coeficientes β_{11} , β_{12} , β_2 , β_3 e β_4 se referem, respectivamente, a: “Preto(a) ou Pardo(a)”, “Indígena ou Amarelo(a)”, “Sexo Masculino”, “Mãe ou a mulher responsável sabe ler/escrever” e “Proficiência em Português”.

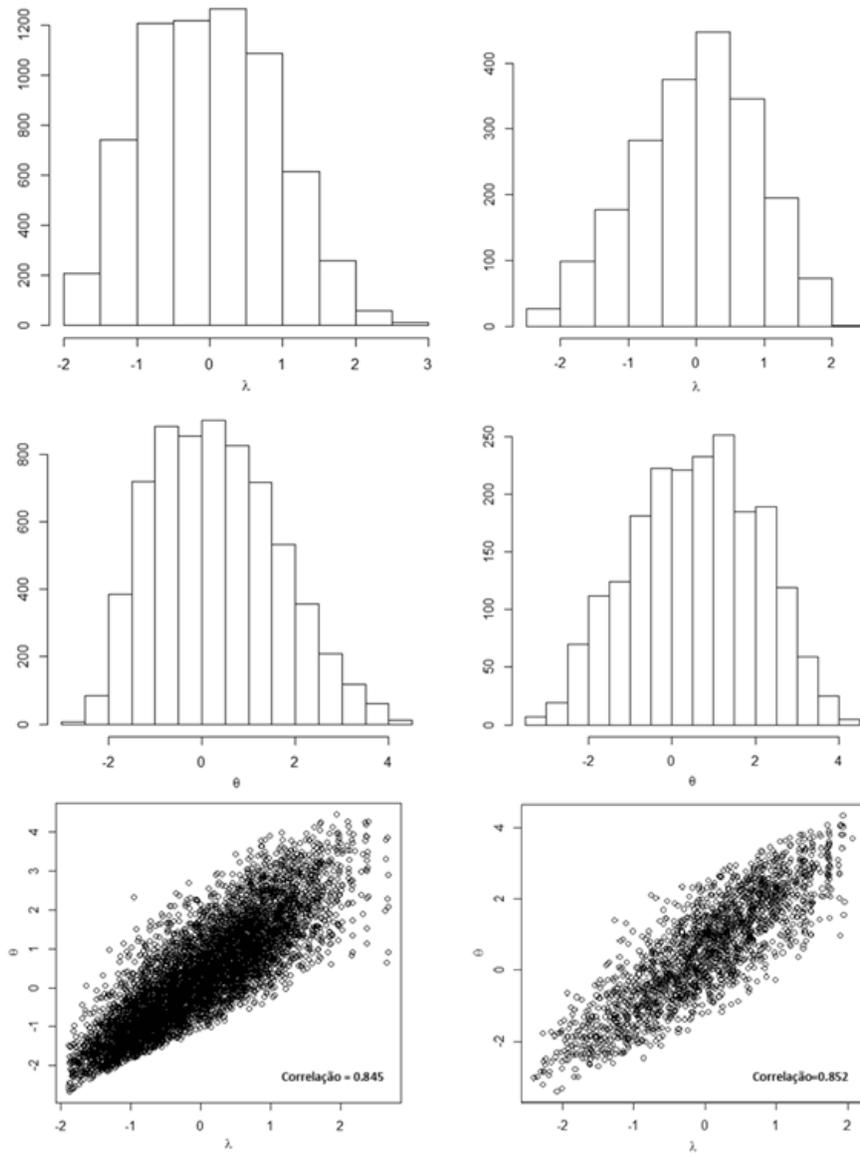
Dados os resultados da análise de regressão, os resultados da tabela 2 indicam uma possível estrutura não aleatória de não resposta associada à segunda pergunta, indicando que alunas do sexo feminino têm maior propensão a não declarar seu sexo, em ambos os estados. A figura 1 mostra as proficiências em Português e Matemática estimadas, bem como a forte relação positiva entre elas.

Tabela 2 Proporções observadas e estimadas de cada resposta com o respectivo número de respostas referentes aos conjuntos de dados do 5º ano EF do Saeb

QUESTÃO	RESPONDIDAS	NÃO RESPONDIDAS	
BA	1	(0.704, 0.083, 0.213) - 5274	(1, 0, 0) - 1394
	2	(0.510, 0.490) - 6275	(0.259, 0.741) - 393
	3	(0.924, 0.076) - 6449	(1, 0) - 219
MG	1	(0.627, 0.066, 0.307) - 1736	(1, 0, 0) - 285
	2	(0.499, 0.501) - 1943	(0.333, 0.667) - 78
	3	(0.974, 0.026) - 1897	(1, 0) - 124

As questões 1, 2 e 3 se referem, respectivamente a: “Como você se considera?”, “Qual o seu sexo?” e “Sua mãe, ou a mulher responsável por você, sabe ler e escrever?”.

Figura 1 Proficiências estimadas em Português (λ) e Matemática (θ) para o 5º ano (Saeb) dos estados da Bahia (esquerda) e de Minas Gerais (direita)



Saeb, 3ª série – Sudeste e Nordeste

Os conjuntos de dados analisados são formados por 741 alunos do Sudeste e 1.148 do Nordeste. Entre os alunos do Sudeste, 22 não responderam a uma das 3 perguntas do questionário, 2 não responderam a duas e 8 não responderam a três, totalizando 50 dados faltantes. Para a região Nordeste, estes números são 58, 3 e 12, totalizando 100 dados faltantes.

O ajuste dos modelos de regressão indica (ver tabela 3), em ambos os estados, uma forte associação positiva entre as proficiências em Português e em Matemática, além de um rendimento médio em matemática um pouco maior de alunos do sexo masculino. Além disso, não ficou evidenciada uma relação entre a proficiência em Matemática e a cor da pele ou o fato de a mãe ou mulher responsável saber ler ou escrever.

Tabela 3 Estimativa dos coeficientes da regressão - média *a posteriori* e intervalo de credibilidade de 99%, referentes aos conjuntos de dados da 3ª Série EM do Saeb – Nordeste (NE) e Sudeste (SE)

COEFICIENTES		ESTIMATIVA
NE	β_{11}	-0.187 (-0.451, 0.083)
	β_{12}	-0.067 (-0.551, 0.401)
	β_2	0.609 (0.366, 0.855)
	β_3	0.319 (-0.109, 0.681)
	β_4	1.376 (1.161, 1.582)
SE	β_{11}	-0.231 (-0.488, 0.029)
	β_{12}	-0.319 (-1.077, 0.360)
	β_2	0.656 (0.394, 0.923)
	β_3	0.256 (-0.526, 0.895)
	β_4	1.396 (1.195, 1.633)

Os coeficientes β_{11} , β_{12} , β_2 , β_3 e β_4 se referem, respectivamente, a: “Preto(a) ou Pardo(a)”, “Indígena ou Amarelo(a)”, “Sexo Masculino”, “Mãe ou a mulher responsável sabe ler/escrever” e “Proficiência em Português”.

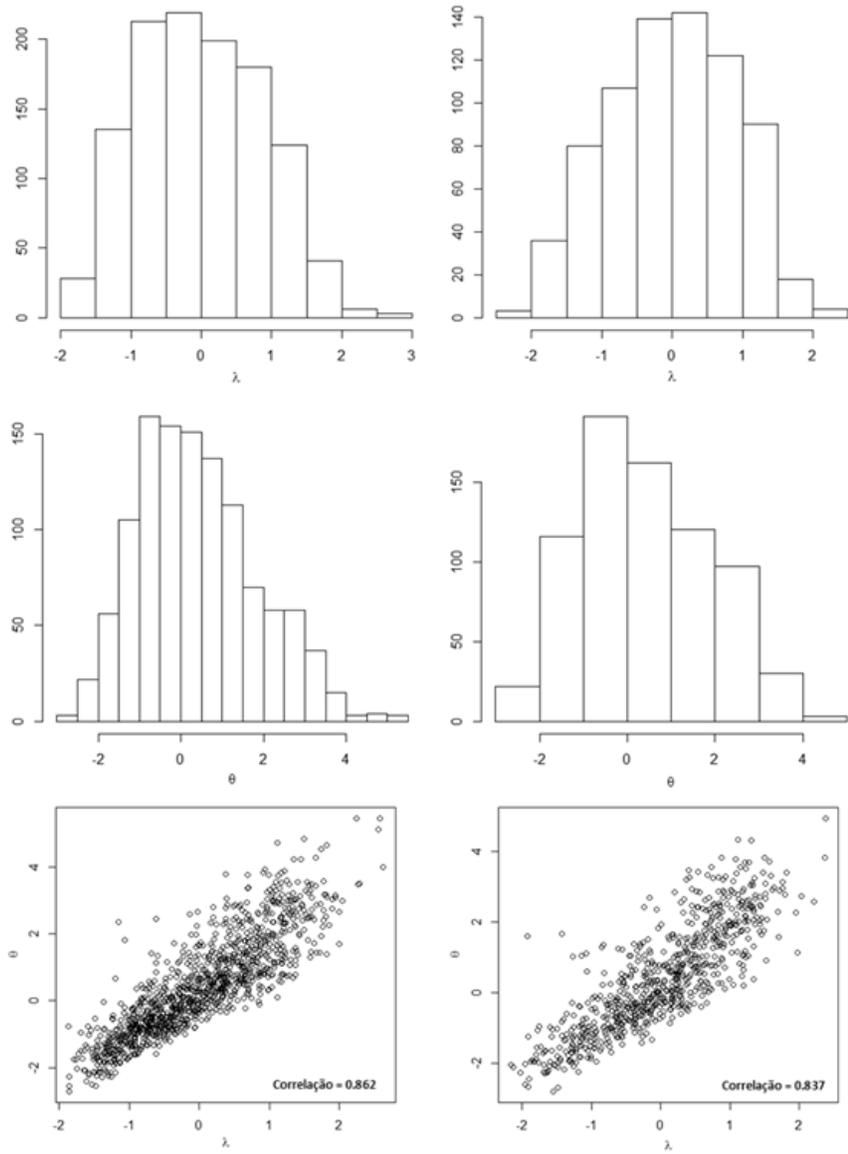
Dados os resultados da análise de regressão, os resultados da tabela 4 indicam uma possível estrutura não aleatória de não resposta associada à segunda pergunta na região Nordeste, indicando que alunas do sexo feminino têm maior propensão a não declarar seu sexo. A figura 2 mostra as proficiências em Português e Matemática estimadas, bem como a forte relação positiva entre elas.

Tabela 4 Proporções observadas e estimadas de cada resposta com o respectivo número de respostas referentes aos conjuntos de dados da 3ª Série EM do Saeb

QUESTÃO	RESPONDIDAS	NÃO RESPONDIDAS
NE	1 (0.664, 0.078, 0.258) - 1100	(1, 0, 0) - 48
	2 (0.431, 0.569) - 1114	(0.206, 0.794) - 34
	3 (0.910, 0.090) - 1130	(1, 0) - 18
SE	1 (0.552, 0.047, 0.401) - 720	(0.857, 0, 0.143) - 21
	2 (0.460, 0.540) - 724	(0.471, 0.529) - 17
	3 (0.973, 0.027) - 729	(1, 0) - 12

As questões 1, 2 e 3 se referem, respectivamente a: “Como você se considera?”, “Qual o seu sexo?” e “Sua mãe, ou a mulher responsável por você, sabe ler e escrever?”.

Figura 2 Proficiências estimadas em Português (λ) e Matemática (θ) para o 5º ano (Saeb) das regiões Nordeste (esquerda) e Sudeste (direita)



Enem – Sudeste e Nordeste

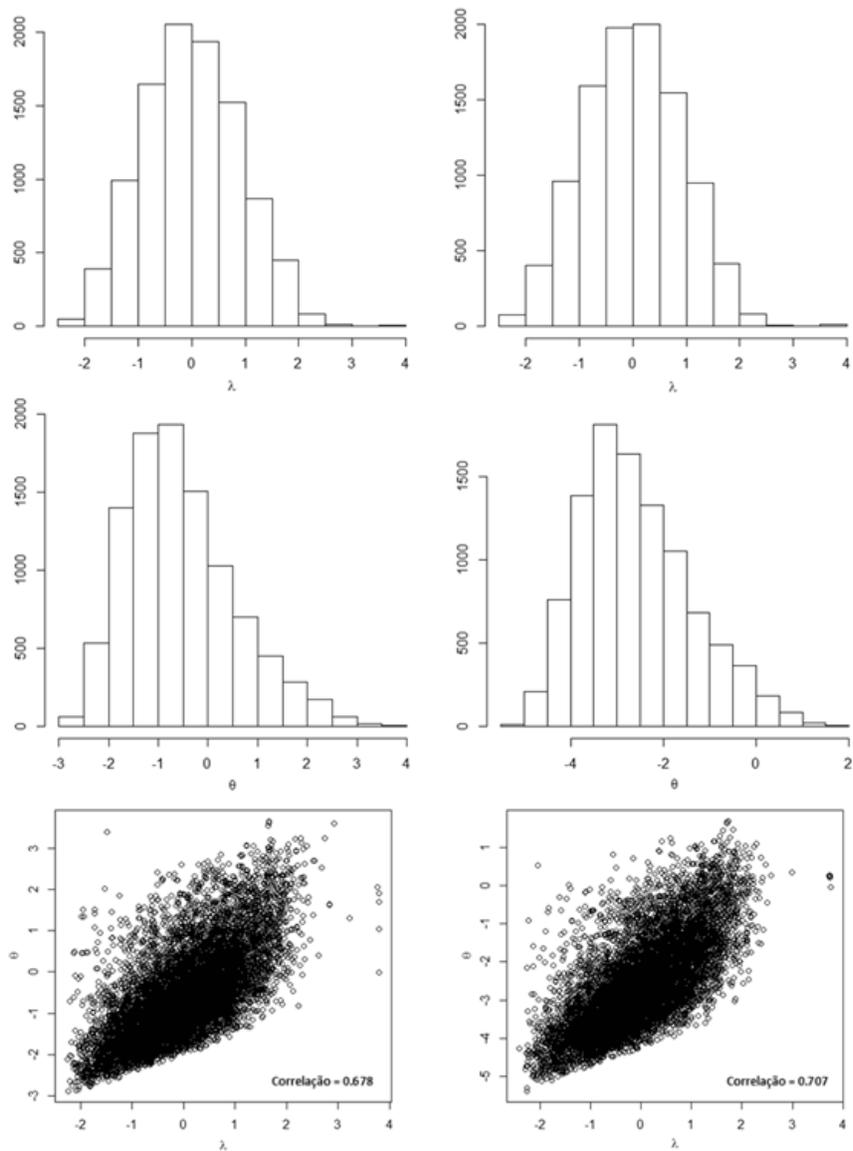
Os ajustes dos modelos de regressão indicam (ver tabela 5), em ambas as regiões, uma forte associação positiva entre as proficiências em Português e em Matemática. Além disso, alunos que estudaram em escolas públicas (sempre ou em parte) apresentaram um rendimento médio em matemática consideravelmente menor que alunos que estudaram somente em escolas particulares. No caso da segunda covariável, alunos que exercem ou já exerceram atividade remunerada tiveram um rendimento médio em Matemática muito maior que os que nunca exerceram na região Sudeste, mas apresentaram um rendimento médio um pouco menor no Nordeste. Por fim, alunos sem acesso à internet em casa apresentaram um rendimento médio um pouco menor do que os que têm acesso, em ambas as regiões. A figura 3 mostra as proficiências em Português e Matemática estimadas, bem como a forte relação positiva entre elas.

Tabela 5 Estimativa dos coeficientes da regressão - média a posteriori e intervalo de credibilidade de 99%, referentes aos conjuntos de dados do Enem – Nordeste (NE) e Sudeste (SE)

COEFICIENTES		ESTIMATIVA
NE	β_{11}	-0.714 (-0.801, -0.625)
	β_{12}	-0.563 (-0.744, -0.373)
	β_2	0.229 (0.161, 0.299)
	β_3	-0.225 (-0.311, -0.143)
	β_4	0.695 (0.648, 0.742)
SE	β_{11}	-0.858 (-0.945, -0.770)
	β_{12}	-0.618 (-0.786, -0.448)
	β_2	-1.834 (-2.022, -1.653)
	β_3	-0.319 (-0.426, -0.220)
	β_4	0.790 (0.743, 0.836)

Os coeficientes β_{11} , β_{12} , β_2 , β_3 e β_4 se referem, respectivamente, a: “Ensino Médio somente em escola pública”, “Ensino Médio parte em pública e parte em privada”, “Não exerce nem nunca exerceu atividade remunerada”, “Não possui acesso à internet na residência”.

Figura 3 Proficiências estimadas em Português (λ) e Matemática (θ) (Enem) das regiões Nordeste (esquerda) e Sudeste (direita)



Discussão

Este artigo analisou conjuntos de dados referentes às duas principais avaliações educacionais no Brasil: Saeb e Enem. O principal objetivo da análise foi identificar possíveis fatores relacionados ao desempenho dos alunos nas respectivas provas de Matemática. Um destes fatores era o desempenho em Português e os demais fatores foram extraídos dos questionários contextuais respondidos aos alunos. No caso do Saeb, foi necessário lidar com uma quantidade considerável de não respostas nos questionários.

A metodologia utilizada consiste em um modelo da TRI que modela a proficiência dos alunos em função dos fatores considerados e trata, de forma adequada, o problema de não respostas ao questionário. Esta metodologia foi proposta em Gonçalves e Dias (2018), em que maiores detalhes sobre a metodologia podem ser encontrados.

Para ambas as avaliações, os resultados indicaram fatores relacionados ao rendimento dos alunos e padrões não aleatórios de não resposta. Em particular, todas as análises indicaram forte relação entre os desempenhos em Português e Matemática. Nas análises do Saeb, alunos do sexo masculino apresentaram melhor rendimento médio e não foi encontrada relação significativa com a cor da pele. Além disso, o fato de a mãe saber ler e escrever parece influenciar positivamente no desempenho de alunos do 5º ano na região Sudeste. Com relação ao Enem, alunos de escolas públicas apresentaram rendimento médio em matemática consideravelmente menor, enquanto alunos que exercem ou já exerceram atividade remunerada tiveram um rendimento médio bem maior no Sudeste e um pouco menor no Nordeste. Finalmente, alunos com acesso à internet tiveram rendimento médio um pouco maior. Com relação aos padrões de respostas, a maioria das análises do Saeb indicou que alunos do sexo feminino têm uma maior propensão a não declarar seu sexo.

As análises mostraram o potencial da metodologia adotada que pode, naturalmente, ser aplicada a variados conjuntos de dados de avaliações educacionais para as quais estejam disponíveis dados de questionários contextuais.

Gonçalves e Dias (2018) ainda propõem generalizações interessantes do modelo utilizado neste artigo. Entre elas, modelos com estruturas hierárquicas para também se considerar questionários aplicados a professores e escolas, seleção de covariáveis e estruturas mais flexíveis para a distribuição das proficiências.

Referências

COLEMAN, J. S. et al. *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office, 1966.

FRANCO, C. et al. O referencial teórico na construção dos questionários contextuais do Saeb 2001. *Estudos em avaliação educacional*, n. 28, p. 39-74, 2003.

GONÇALVES, F. B.; DIAS, B. C. C. Bayesian modelling of the abilities in dichotomous IRT models via regression with missing values in the covariates. A ser publicado no *Brazilian Journal of Probability and Statistics*, 2018.

JENCKS, C. et al. *Inequality: a reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.

LAROS, J. A.; Marciano J. L. Índices educacionais associados à proficiência em língua portuguesa: um estudo multinível. *Avaliação Psicológica*. V. 7, n. 3, 2008.

Flávio Bambirra Gonçalves

Doutor em Estatística pela University of Warwick.
Professor da UFMG - Belo Horizonte (MG), Brasil;
fbgoncalves@est.ufmg.br

Bárbara da Costa Campos Dias

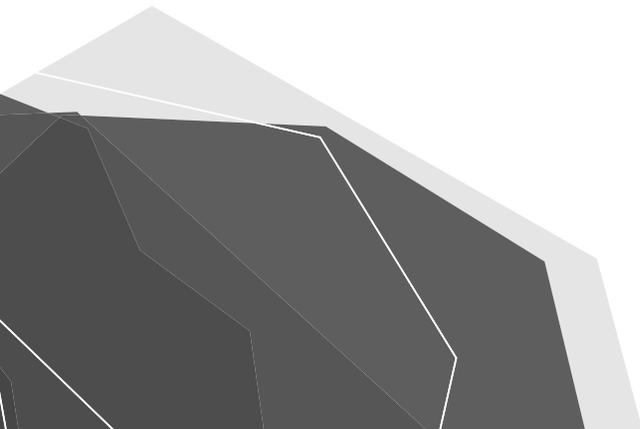
Mestrado em estatística na Universidade Federal de Minas Gerais, Belo Horizonte (MG), Brasil;
bccd89@gmail.com

Artigo submetido em 25/10/2017

Aprovado em 12/06/2018

**RE
SE
NHA**

review



PADRÕES PARA EDUCAÇÃO E PSICOLOGIA

STANDARDS FOR EDUCATIONAL AND PSYCHOLOGICAL TESTING

Cristiane Faiad

Carlos Manoel Lopes Rodrigues

AERA, APA, & NMCE. Standards for Educational And Psychological Testing.

O *Standards*, traduzido como Padrões para Educação e Psicologia, foi publicado pela *American Educational Research Association* (AERA), *American Psychological Association* (APA) e pela *National Council on Measurement in Education* (NCME) como fonte de orientação para construção e uso de testes, na tentativa de promover melhores práticas no uso da medida, subsidiar a qualificação dos mesmos e orientar a busca por critérios psicométricos para adequada interpretação dos escores fornecidos pelos instrumentos. Documentos desta natureza são propostos desde 1954, por diferentes comitês de gerenciamento de normas, embora a versão no padrão do *Standards* tenha sido desenhada em 1999. Já, em 2014, houve a necessidade de que o documento fosse atualizado, sendo esta a última versão disponível. Uma importante consideração é feita no *Standards*, ao explicar que, apesar de diferenças no uso do termo escalas, inventários e testes, este último é utilizado para se referir a todo tipo de instrumento avaliativo. O teste é definido como um procedimento que permite a avaliação de uma amostra de comportamento de um examinando, em determinado domínio, por meio de um processo padronizado. Espera-se, contudo, que o documento seja aplicado para além das medidas utilizadas nos campos da educação e da psicologia, como habilidade, aptidão, realização, atitudes, interesses, personalidade, funcionamento cognitivo e saúde mental, estendendo-se para outras medidas no campo da avaliação. Enquanto documento que estabelece diretrizes, chama a atenção de que contempla orientações importantes e mínimas na construção de medidas, embora não isente olhares mais técnicos, quando se trata de contextos judiciais e de normativas existentes sobre o assunto.

A primeira parte da obra se concentra nas questões relativas aos fundamentos do desenvolvimento e uso dos testes, englobando temas como validade, confiabilidade e equidade. Quanto à validade, o *Standards* apresenta sua importância e uma definição de que “refere-se ao grau em que a evidência e a teoria suportam a interpretação dos escores do teste para o uso proposto pela medida” (p. 9), ou seja, considerando um processo de acúmulo de evidências relevantes que proporcionem uma base científica sólida para a interpretação dos resultados dos testes. Esse é um processo contínuo de responsabilidade de desenvolvedores e usuários dos testes, que se inicia com o planejamento do instrumento e continua por meio de seu desenvolvimento e implementação, envolvendo cuidados com os possíveis vieses que possam emergir por representações inadequadas do construto, ou por aspectos do instrumento, tais como formato do teste, condições de administração, e nível da linguagem. As evidências de validade são organizadas no *Standards* como: a) orientadas ao conteúdo; b) relativas ao processo cognitivo subjacente e ao processo de resposta; c) relativas à estrutura interna e às relações com construtos conceitualmente; d) relacionadas aos critérios; e e) como evidências baseadas nas consequências do teste. O destaque dado para as consequências da testagem alinha-se à preocupação presente em toda a obra quanto à responsabilidade e ao impacto social que o processo de avaliação acarreta.

Quanto à confiabilidade e precisão, parte-se da noção geral de consistência dos resultados através das várias instâncias dos procedimentos de teste, sendo a necessidade de precisão proporcional à importância da interpretação e das decisões decorrentes dos resultados dos testes. São apresentadas formas de análise de confiabilidade e precisão (como, por exemplo, o erro-padrão de medida, a Teoria da Generalizabilidade e a Teoria da Resposta ao Item), mas na prática a avaliação da confiabilidade/precisão é feita utilizando-se diferentes coeficientes, a depender do foco da análise e do modelo de medida em uso. O destaque é dado, na primeira parte, à equidade nos testes, com a proposição de dois novos conceitos: o de acessibilidade e o de *design* universal. Por acessibilidade, entende-se que todos os examinados devem ter oportunidade de realizar o teste, levando em consideração sua posição/condição específica; e o *design* universal compreendido como uma abordagem que busca maximizar a acessibilidade para todos os candidatos que se pretende medir. Especificamente o *design* universal requer construtos precisamente definidos, procedimentos e instruções simples, claros e intuitivos e máxima legibilidade e compreensibilidade. O atendimento deste requisito inclui uma forma

mais globalizante de pensar a construção de instrumentos, considerando questões técnicas e sociais mais amplas.

A segunda parte do *Standards* denominada “operações” compreende recomendações e procedimentos desde o projeto e desenvolvimento do teste, até o desenvolvimento de normas e padrões de interpretação. Esta parte inclui, também, duas sessões específicas quanto aos direitos e deveres dos desenvolvedores e usuários dos testes. Além dos procedimentos e técnicas já reconhecidos, alguns pontos são destacados, como a necessidade de envolver psicometristas na fase de desenvolvimento, avaliar a relação custo-benefício e a utilidade das ações de desenvolvimento dos testes. O próximo destaque é dado à criação de normas e padrões de interpretação dos testes, com a discussão centrada nas formas de redução de vieses de interpretação em prol de decisões com maior generalidade e equidade. Esta preocupação é reforçada no detalhamento dado à definição das formas de aplicação dos testes e no fornecimento de informações detalhadas sobre o teste para o futuro uso adequado. O apontamento dos direitos e deveres dos desenvolvedores e usuários dos testes ressalta essa preocupação ética que perpassa todo o *Standards*.

Por fim, a terceira parte é direcionada à aplicação dos testes em diferentes contextos, sendo destacado o uso no psicodiagnóstico (capítulo 10), no contexto organizacional e de acreditação profissional (capítulo 11), nas avaliações educacionais (capítulo 12) e na avaliação de programas e políticas públicas.

Quando comparada a versão atual com a anterior, de 1999, observa-se que os fundamentos e operações foram destacados para as duas primeiras partes da obra, dedicando assim maior espaço para estas questões. O uso de capítulos sequenciais específicos sobre responsabilidade e direitos de desenvolvedores e usuários de testes também representa uma mudança introduzida nesta versão, já que na versão de 1999 estes tópicos vinham em partes separadas da obra. A discussão sobre equidade, que, na versão antecessora, ocupava a segunda parte da obra, agora permeia todos os capítulos como um fio condutor que alinha toda a preocupação ética e social das organizações autoras. Assim, a versão atual do *Standards* apresenta, além do conjunto de boas práticas na construção e uso de instrumentos de medida, um retrato do desenvolvimento da área ao incorporar inovações técnicas e questões éticas relevantes para o campo. A obra, portanto, não é só uma ótima referência para o processo de construção de medidas, mas também para o usuário final dos instrumentos.

Cristiane Faiad

Doutora em Psicologia Social, do Trabalho e das Organizações (PSTO) pela Universidade de Brasília.

Professora da Universidade de Brasília, Brasília (DF), Brasil;
crisfaiad@gmail.com

Carlos Manoel Lopes Rodrigues

Mestre em Psicologia Social, do Trabalho e das Organizações pela Universidade de Brasília.

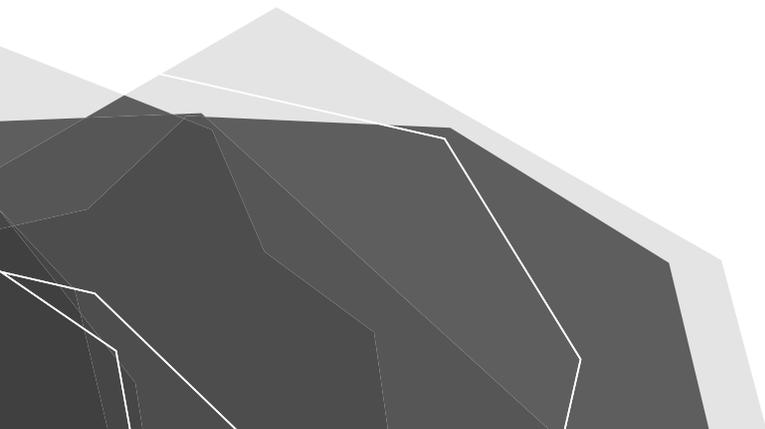
Professor Assistente do Centro Universitário de Brasília – UNICEUB, Brasília (DF), Brasil;
prof.carlos.manoel@gmail.com

Resenha submetida em 30/05/2018

Aprovada em 11/06/2018

**EN
TRE
VIS
TA**

interview



O PAPEL DA PSICOMETRIA NA TESTAGEM EM EDUCAÇÃO E PSICOLOGIA

THE ROLE OF PSYCHOMETRY IN THE TEST IN EDUCATION AND
PSYCHOLOGY

EL PAPEL DE LA PSICOMETRÍA EN EL ENSAYO DE EDUCACIÓN Y
PSICOLOGÍA

Professor Luiz Pasquali

Professor Pesquisador Associado da Universidade de Brasília. Coordenador do Laboratório de Pesquisas em Avaliação e Medidas. Presidente do Instituto Movens, Sócio da LabPAM Saber e Tecnologia. Doutor em Psicologia pela Faculté de Psychologie et des Sciences de l'éducat Université Catholique de Louvain, Bélgica.

Examen – O senhor é muito conhecido entre os estudiosos, principalmente na área de psicometria, pode nos contar como foi a sua trajetória?

Luiz – Eu estudei em um seminário. Depois de um tempo, me mandaram para Roma, mas fiquei meio decepcionado e fui para a Bélgica, para estudar Psicologia. Escolhi essa área porque já tinha formação, aqui no Brasil, em Petrópolis, na área de Pedagogia.

Nos anos 1960, como vivia sozinho, ia para os Estados Unidos todos os anos para ganhar um dinheirinho para pagar a universidade. Trabalhei até defender a tese, em 1970. Coletei os dados para a pesquisa em Rochester, Nova York, no colégio que me ofereceram para trabalhar sobre a simbolização do conceito de Deus a partir das imagens parentais. Também fiz toda a análise lá, já que tive que trabalhar com análise fatorial. Usava computadores e, na época, anos 70, não havia computadores manuais, então trabalhava na IBM mesmo.

Da Bélgica, imigrei para os EUA e casei com uma americana, trabalhei durante três anos em faculdades e departamentos de Psicologia, perto de Kalamazoo, na região de Michigan, região muito fria. Fiquei lá quatro anos, mas eu sempre pensava em voltar para o Brasil, para trazer mais um PhD. Finalmente, surgiu a chance: a PUC de Porto Alegre me ofereceu uma vaga. Eles falaram que tinha pós-graduação, clínica especializada, mas eu atuei no reconhecimento do mestrado. Só que foi uma decepção atrás da outra e eu quis sair de lá.

Em dado momento, conheci o Luiz Hesketh, então chefe do departamento de Psicologia da UnB, que insistiu para que eu viesse para Brasília. Depois de um ano, me chamaram para a capital, em 1975. Estou aqui desde então. Me chamaram porque queriam montar um mestrado também. Havia um mestrado na área de Psicologia Social e havia uma pós-graduação sobre a teoria do behaviorismo. Então, precisavam de um pessoal para verificar no MEC a formação do mestrado. E esse foi o começo da minha história.

Examen – Qual a sua opinião sobre a contribuição da Psicometria para a avaliação educacional?

Luiz – A Psicometria existe para produzir parâmetros científicos na área da avaliação. Ela procura substituir o subjetivismo por critérios mais objetivos. Objetivar, na Psicometria, significa praticamente quantificar as coisas. A Psicometria usa um modelo matemático que introduz exatidão na conversa, para explicar o comportamento humano. Você tem uma modelagem matemática em que as coisas são mais exatas e precisas e em que o subjetivismo não tem nada a ver.

Examen – Nesse sentido, ela traz mais objetividade para as avaliações educacionais?

Luiz – Essa é a ideia: substituir a conversa, a linguagem, a semântica pelo número. O número tem muito mais precisão, ele não tem conotações. A semântica tem muitas conotações por estar

dentro de uma cultura que vem de séculos. Cada palavra tem um significado, dependendo da cabeça de quem a está usando. O número não tem um significado preciso – embora, na Psicometria, não seja o número matemático, onde o 1 é 1, o 2 é 2, mas sim um número estatístico. É um número um pouco misturado, vamos dizer assim. O 1 pode até ser 2, dentro de uma faixa. Mas, mesmo assim, a gente pode definir o tamanho desse intervalo. Aposto que não dá para definir o tamanho da semântica de uma palavra. Depende da cultura, que varia muito. Psicometria é introduzir precisão, exatidão, e, com isso, você pode modelar coisas, simular coisas. Com palavras, não dá para simular.

Examen – Atualmente, temos um padrão bom na adoção da psicometria nas avaliações educacionais, mas precisamos melhorar, principalmente na fase pré-prova. Como o senhor avalia esse processo?

Luiz – Na Psicologia, existe uma oposição bastante grande à mentalidade da psicometria. No Brasil, por exemplo, o meio da Psicologia não é a Psicometria, é a Psicanálise e essas coisas aí. Ainda assim, têm os geridos da Psicometria, que são os pedagogos. Pedagogo, realmente, não quer saber dessa conversa. Eles gostam muito de filosofia, de fundamentos – o que é importante, sim, mas fazer progredir a ciência dentro dessa área não é por aí. Psicometria é um instrumento importantíssimo para avançar um conhecimento realista, objetivo. Não é uma mágica. É um instrumento valioso para o negócio, sem que, com isso, a gente dê menos importância à teoria. A teoria é fundamental, ótima para produzir hipóteses, mas que precisa ser testada empiricamente, e ser testada empiricamente, e aí onde entra a Psicometria como instrumento de grande valia.

Examen – Como o senhor vê a questão de validade na área de Educação?

Luiz – A Psicologia trabalha com construtos, a Pedagogia também. São construtos diferentes, mas o processo de avaliar esses construtos é o mesmo. A psicometria é idêntica na Pedagogia, na Psicologia e em ciências psicossociais, de modo geral. Todo aquele que trabalha com o construto ou dentro

da visão da TRI (Teoria de Resposta ao Item) trabalha com o teta — escore representado pela oitava letra do alfabeto grego —, e, portanto, precisa trabalhar Psicometria. Só que esse teta, seja pedagógico, seja psicológico, é sempre por processos mentais. Tem que usar Psicometria para tudo isso. E qual é a validade? No fundo, a validade é você verificar a correspondência, a congruência, a relevância, a relação que existe entre o comportamento e o processo mental em jogo, o teta.

Examen – Como as investigações de evidências de validade poderiam melhorar as medidas educacionais?

Luiz – No momento, nós não temos outra coisa para fazer. Precisamos trabalhar com as cinco evidências de validade. O Satepsi (Sistema de Avaliação de Testes Psicológicos) não aceita um teste que tenha demonstração de evidência de validade de um só tipo. Tem que ter, pelo menos, dois tipos. Por enquanto, a demonstração considera as cinco evidências de validade. Então, como evidências circunstanciais: duas circunstâncias já fundamentam, pelo menos, um teste, e dizem que um instrumento é válido. Além da verificação da estrutura interna dos instrumentos, tem que fazer um levantamento do que está na literatura, no nosso Google, na nossa cultura científica sobre a história e desenvolver instrumentos. A validade é mais que isso. Você vê que, no fundo, nós perdemos o conceito de validade. Porque não é mais o teste que é válido, mas são as consequências e decisões de um teste que nós tomamos sobre o comportamento da pessoa. É possível prejudicar uma pessoa baseado nessa informação.

Examen – E a validade de construto, então, seria o quê?

Luiz – É, no momento, o levantamento na internet – quando falo de internet, é o consenso – dos pesquisadores da área sobre um construto. Por exemplo, ansiedade. O que é? Não existe consenso. Tem o psicanalista que pensa de um modo, o behaviorista, de outro modo, mas existe uma grande literatura sobre o tema. Então, você faz como nas teses de mestrado e doutorado: “Adão e Eva disseram isso. Outro disse isso”. No fim, você vai ter que

pôr a sua conclusão. Você, que vai construir o teste, diz “eu entendo ansiedade assim. Por quê? Porque eu acho que é assim.” A fundamentação teórica de um teste é fundamental. É preciso levantar a história sobre o construto, o que existe na literatura científica. Isso é obrigação. Se não fizer, é avacalhação, não é ciência. Fica mais fácil quando se tem uma ideia ou hipótese do que se pode fazer. Para que inventar uma hipótese nova e não levar em conta o que já existe? É preciso fazer o levantamento nem que seja para refutar a hipótese.

Examen – Para o senhor, o que é validade?

Luiz – A validade é um conjunto de comportamentos válidos, no sentido de que eu posso tirar conclusões sobre um sujeito a partir de um teste, sobre uma possível habilidade existente, o teta, se algumas considerações sobre o instrumento forem levadas em conta, como o construto e o conteúdo, eu não me torno arbitrário ao tomar decisões sobre o sujeito. Se ele me dá as garantias, então, posso tirar as conclusões e tomar decisões sobre seu comportamento. Assim, não é o instrumento que é válido, mas ele possui características que me dão certa segurança para tomar decisões justas. Validade é isso, não tem outra forma de vê-la. Não existe validade do teste, existem evidências de que esse instrumento produz informação válida. Informação séria, segura, aceitável. Isso é a validade de um instrumento.

