

Análise transversal múltipla dos indicadores de qualidade derivados do ENADE

HELIO RADKE BITTENCOURT^I

MARION CREUTZBERG^{II}

JULIO BERTOLIN^{III}

<http://dx.doi.org/10.22347/2175-2753v15i48.3998>

Resumo

A realização de estudos de análise dos resultados derivados do Sistema Nacional de Avaliação da Educação Superior (SINAES) pode contribuir para a avaliação da qualidade da educação superior brasileira. Todavia, a literatura especializada apresenta uma carência de estudos transversais múltiplos e longitudinais. Nesse sentido, este estudo tem por objetivo realizar uma análise comparativa dos resultados do Enade ao longo do tempo no que se refere aos indicadores de qualidade. O estudo levou em conta as restrições impostas por questões metodológicas. Os resultados evidenciam que o indicador institucional apresentou um pequeno aumento nos últimos cinco anos causado, principalmente, pela melhora nos componentes do conceito preliminar de curso (CPC) relacionados à titulação de mestrado e regime de trabalho dos docentes que sofrem efeito de truncamento. O estudo também conclui que os indicadores são inadequados para avaliação da evolução ou involução do sistema ao longo do tempo.

Palavras-chave: Educação superior; Avaliação educacional; Exame Nacional de Desempenho dos Estudantes.

Submetido em: 06/07/2022

Aprovado em: 12/09/2023

^I Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0002-7018-6926>; e-mail: heliorb@pucrs.br.

^{II} Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre (RS), Brasil; <http://orcid.org/0000-0001-8366-3560>; e-mail: marionc@pucrs.br.

^{III} Universidade de Passo Fundo (UPF), Passo Fundo (RS), Brasil; <http://orcid.org/0000-0002-5547-1550>; e-mail: julio@upf.br.

Multiple cross-sectional analysis of quality indicators extracted from ENADE

Abstract

The studies of the results derived from the National System for the Evaluation of Higher Education (SINAES) can contribute to the evaluation of the quality of Brazilian higher education. However, the literature shows a lack of multiple cross-sectional and longitudinal studies. This study aims to compare the results of the large-scale exam over time with regard to quality indicators. The study took into account the restrictions imposed by methodological issues. The results show that the institutional indicator showed a small increase in the last five years, mainly caused by the improvement in two components of the preliminary course concept (CPC) - master's degrees and the work regime of faculty staff - due to truncation effect. The study also concludes that the indicators are inadequate for evaluating the evolution or involution of the system over time.

Keywords: Higher education; Educational assessment; Large-scale evaluation.

Análisis transversal múltiple de indicadores de calidad derivados del ENADE

Resumen

Estudios de los resultados del Sistema Nacional de Evaluación de la Educación Superior (SINAES) pueden contribuir acerca de la evaluación de la calidad de la educación superior brasileña. Sin embargo, la literatura muestra una falta de estudios transversales múltiples y longitudinales. Este estudio tiene como objetivo comparar los resultados del examen a gran escala a lo largo del tiempo con respecto a los indicadores de calidad. El estudio tuvo en cuenta las restricciones impuestas por cuestiones metodológicas. Los resultados muestran que el indicador institucional mostró un pequeño aumento en los últimos cinco años, principalmente debido al incremento en dos componentes del concepto de curso preliminar (CPC) – la titulación de maestría y el régimen de trabajo del personal docente – debido al efecto de truncamiento. El estudio también concluye que los indicadores son inadecuados para evaluar la evolución o involución del sistema en el tiempo.

Palabras clave: Educación superior; Evaluación de la educación; Evaluación a gran escala.

Introdução

A Lei 10.861, de 14 de abril de 2004 (BRASIL, 2004), instituiu o Sistema Nacional de Avaliação da Educação Superior (SINAES) tendo como finalidades a melhoria da qualidade da educação superior, a orientação da expansão da sua oferta, o aumento permanente da sua eficácia institucional e efetividade acadêmica e social e, especialmente, a promoção do aprofundamento dos compromissos e responsabilidades sociais das instituições de educação superior.

Dentre essa diversidade de relevantes objetivos, o controle e indução da melhoria da qualidade das instituições e dos cursos de graduação vêm assumindo um papel de centralidade. A grande expansão quantitativa ocorrida nas últimas duas décadas, em larga medida baseada na modalidade a distância, passando de menos de 3 milhões para mais de oito milhões de estudantes matriculados e de pouco mais de 1.000 para aproximadamente 2.500 instituições (BRASIL, 2022), certamente tensionou a dimensão da qualidade do sistema. Além disso, com a democratização do acesso por meio, especialmente, da Lei de Cotas e do Prouni, o perfil e o nível socioeconômico médio dos matriculados está mudando, impactando na aprendizagem e desempenho dos estudantes (MCCOWAN; BERTOLIN, 2020).

Assim, neste contexto de importantes transformações, a avaliação do comportamento qualitativo do sistema como um todo, com julgamento de melhora ou piora no período do SINAES, vem se constituindo como um grande desafio do sistema de avaliação.

Embora a proposta original do SINAES contemple três dimensões – avaliação institucional, avaliação de cursos e avaliação do desempenho dos estudantes, mais conhecida como Enade – esta última tem sido, sem dúvida, a componente de maior notoriedade do sistema (BITTENCOURT; CASARTELLI; RODRIGUES, 2010; RISTOFF; GIOLO, 2006). A divulgação simultânea dos resultados ocasiona imediata repercussão dos resultados nas diferentes mídias e tem sido utilizada para a construção de *rankings*. De fato, o conjunto de indicadores derivados do Enade possibilita um amplo diagnóstico da qualidade das instituições avaliadas, contudo, análises superficiais dos resultados podem ocultar importantes aspectos que dependem da metodologia e da devida contextualização.

A realização de estudos de análise dos resultados derivados do Enade pode contribuir muito para avaliação do SINAES, visto que os mesmos carregam informações valiosas sobre os estudantes, cursos e suas respectivas instituições. Tais

informações possibilitam que o desempenho seja comparado de acordo com o perfil do aluno, curso, organização acadêmica, categoria administrativa, região geográfica, entre outros.

As 16 edições já realizadas do Enade (de 2004 a 2019) disponibilizam um grande conjunto de dados, contudo, a literatura especializada apresenta uma carência de estudos transversais múltiplos e longitudinais, bem como de trabalhos derivados dos microdados do Enade. Há uma variedade de estudos desenvolvidos sobre o processo de avaliação, seus aspectos técnicos e concepções teóricas, dentre os quais se pode destacar alguns trabalhos: MARCHELLI, 2007; BITTENCOURT; VIALI; CASARTELLI, 2008; POLIDORI, 2009; BARREYRO; ROTHEN, 2014; VERHINE, 2015.

Em estudo recente, Haar (2021) aponta erros no cálculo do IDD a partir da edição de 2014 do Enade e apresenta resultados comparativos entre os valores divulgados pelo INEP e os calculados por meio do modelo correto. Outros estudos também elencam pontos positivos e negativos relativos ao sistema, acompanhados de proposições de melhorias. Verhine (2015) discute os conceitos de avaliação e regulação, destacando a dificuldade de vincular a regulação a partir de indicadores que simplificam realidades complexas. A mesma preocupação já fora levantada por Barreyro (2008).

No que se refere a estudos longitudinais, tema central deste estudo, deve-se mencionar que alguns dos trabalhos publicados nos Seminários Regionais do INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2015) tratam da trajetória das Comissões Próprias de Avaliação no período de implantação do SINAES, o que confere um caráter longitudinal aos mesmos. Dentre os poucos estudos transversais múltiplos realizados fora do âmbito do INEP, destacam-se três: a) estudo de abrangência institucional (GURGEL, 2010); b) estudo realizado em âmbito estadual (CORTELAZZO; RIBEIRO, 2013); c) abordagem longitudinal no período de 2006 a 2012 dos resultados do Enade realizada por pesquisadores brasileiros e portugueses considerando cursos da área de Economia e Negócios (PEREIRA; MACHADO-TAYLOR; ARAUJO; TEIXEIRA, 2017). Com base na revisão da literatura, pode-se afirmar que a avaliação dos resultados dos cursos do 1º ciclo nos anos de 2004 e 2007 (BRASIL, 2011), constitui o trabalho desse tipo mais abrangente realizado no âmbito do SINAES até o momento.

Este estudo tem por objetivo apresentar uma análise comparativa dos resultados do Enade divulgados em todas as edições, desde 2007, no que se refere

aos indicadores de qualidade de curso (Enade, IDD, CPC) e o indicador institucional (IGC). Há limitações impostas pela metodologia de cálculo dos indicadores que dificultam a comparação temporal dos resultados, sendo a principal delas o uso do afastamento padronizado por área de avaliação e edição. Outra limitação se deve às mudanças na composição e nos pesos dos componentes de cada indicador, o que impediu a realização de algumas comparações.

Um breve histórico das mudanças nos indicadores

O período compreendido entre 2004 e 2018 foi marcado por modificações nos indicadores derivados do Enade, tanto na forma de cálculo como na divulgação dos resultados. Conforme Griboski (2012), a concepção do Enade parte da premissa de que ele seja um indutor da qualidade na educação superior brasileira. Assim como o exame promove melhorias no sistema, o próprio exame se retroalimenta visando o aperfeiçoamento do processo avaliativo. O Quadro 1 resume as principais modificações ocorridas no processo avaliativo e nas ponderações do CPC.

Quadro 1 – Esquema das principais modificações ocorridas no Enade no período de 2004 até 2018

		Edição do Enade															
Descrição das modificações		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	
Participação	Amostrai																
	Censitária																
Estágio do curso	Ingressantes + Concluintes																
	Concluintes																
Indicadores	Enade e IDD																
	CPC e IGC																
Cálculo do IDD	Nota dos ingressantes																
	ENEM do grupo																
	ENEM individualizado																
Percepção discente	Uso de duas questões do QE																
	Instrumento psicométrico																
Pesos no CPC	Enade				40%		30%		20%					20%			
	IDD				30%		30%		35%					35%			
	Instalações e Infraestrutura				3%		5%		7,5%					5%			
	Organização didático-pedagógica				8%		5%		7,5%					7,5%			
	Oportunidades de Ampliação				0%		0%		0%					2,5%			
	Professores com doutorado				12%		20%		15%					15%			
	Professores com mestrado				0%		5%		7,5%					7,5%			
	Professores em regime parcial / integral				7%		5%		7,5%					7,5%			

Fonte: Os autores (2022).

O Quadro 1 revela que houve mudanças quanto à participação. De início, apenas parte dos estudantes ingressantes e concluintes eram selecionados para a realização da prova, na forma de amostra, o que passou a ser censitário a partir de

2009. Na sequência, quando o ENEM se tornou a referência de desempenho de entrada, os ingressantes, ainda que inscritos no ENADE, foram dispensados da prova e do questionário, permanecendo apenas os estudantes concluintes dos cursos, no respectivo ciclo.

Após o 1º ciclo avaliativo ter sido completado (2004-2006) e no bojo de discussão acerca da viabilidade do SINAES, que previa a avaliação *in loco* sistemática para todos os cursos de graduação e instituições, foram instituídos os indicadores CPC e IGC, por meio da republicação da Portaria Normativa nº 40 (BRASIL, 2010). No âmbito dos cursos de graduação, o CPC se tornou uma medida de avaliação que impactou em processos regulatórios, determinando a dispensa de realização de avaliação *in loco* para a Renovação de Reconhecimento de cursos, mantendo-a apenas em cursos com conceitos inferiores a 3, para os atos de Autorização e de Reconhecimento e para os casos de Supervisão. O questionário do estudante foi reformulado e, a partir de 2013, todas as suas questões foram utilizadas no cálculo. Os pesos atribuídos aos diferentes componentes foram modificados ao longo dos ciclos, com destaque à diminuição no peso da prova. O IGC igualmente foi aplicado como parâmetro para processos regulatórios, de IES e de Curso, conforme instituiu a referida Portaria.

A existência de uma longa série histórica dos indicadores pode ser passível de uma análise simplificada, a qual ocultaria parte das reais causas de variação.

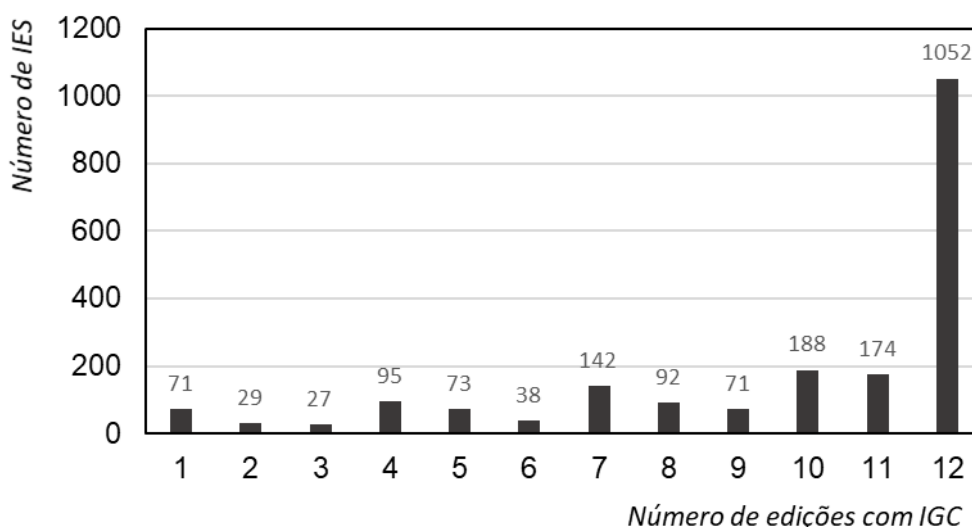
Na próxima seção, é explanado o método utilizado neste estudo visando contornar as limitações impostas pelas alterações nos pesos e pelo uso do afastamento padronizado no cálculo dos indicadores. Ao longo da seção de Resultados também são apresentados alguns detalhamentos necessários para realização das comparações.

Método

O ponto de partida para a realização da análise transversal múltipla foi a organização de todas as planilhas de resultados disponibilizadas na página do INEP (BRASIL, 2021). As últimas edições com divulgação completa de resultados serviram de base para as comparações: no caso do IGC, a última edição foi 2018 e do CPC, foi referente ao triênio 2016-17-18. Isso quer dizer que foi feita uma análise retrospectiva das instituições de ensino superior (IES) avaliadas nestas edições a partir da listagem de 2018.

Embora no ano de 2018 um total de 2.052 instituições tenham recebido um conceito IGC, ao retrocedermos até o ano de 2007, percebe-se que pouco mais da metade das IES (1.052 ou 51,3%) tiveram a atribuição de um conceito IGC em todas as doze edições (Figura 1).

Figura 1 – Distribuição do número de IES conforme o número de edições com atribuição de conceito IGC



Fonte: Os autores (2022) adaptado de BRASIL (2021).

Esse grupo de 1.052 IES foi considerado para fins de comparação do IGC. Há, contudo, as mudanças metodológicas ocorridas que não permitem uma comparação dos resultados durante todo o período de existência do IGC. Constatou-se que, desde 2014, não houve modificações no cálculo do indicador institucional, de forma que o período de análise neste artigo para fins de IGC ficou compreendido entre 2014 e 2018. Houve, ainda, a necessidade de exclusão de 113 instituições que mudaram de organização acadêmica durante esse período, a maior parte delas (92%) de faculdades que se tornaram centros universitários. Portanto, tem-se um conjunto de 939 instituições que tiveram IGC em todas as edições e não tiveram mudanças de organização acadêmica no período de 2014 até 2018.

Em relação ao CPC e seus respectivos insumos, cabe ressaltar a organização por ciclos avaliativos trienais (Quadro 2). De 2007 até 2018, ocorreram quatro edições de cada um dos grupos de cursos que compõem os ciclos avaliativos. A partir de 2019, houve mudanças na configuração dos ciclos, motivo pelo qual o ano de 2019 não foi incluído neste estudo.

Quadro 2 – Configuração dos ciclos avaliativos do Enade conforme as principais áreas avaliadas

Ciclo	Anos	Principais áreas avaliadas nos ciclos
I	2007 – 2010 – 2013 – 2016	Agronomia, Biomedicina, Educação Física (B), Enfermagem, Farmácia, Fisioterapia, Fonoaudiologia, Medicina, Medicina Veterinária, Nutrição, Odontologia, Zootecnia, Tecnológicos relacionados às áreas.
II	2008 – 2011 – 2014 – 2017	Arquitetura e Urbanismo, Artes Visuais, Ciência da Computação, Ciências Biológicas, Ciências Sociais, Educação Física (L), Engenharias, Filosofia, Física, Geografia, História, Letras, Matemática, Música, Pedagogia, Química, Sistemas de Informação, Tecnológicos relacionados às áreas.
III	2009 – 2012 – 2015 – 2018	Administração, Administração Pública, Ciências Contábeis, Ciências Econômicas, Design, Direito, Jornalismo, Psicologia, Publicidade e Propaganda, Relações Internacionais, Secretariado Executivo, Serviço Social, Teologia, Turismo, Tecnológicos relacionados às áreas.

Fonte: Os autores (2022).

Para formar a base de cursos a ser comparada, foi necessário encontrá-los nas planilhas pretéritas. De 2016 em diante, o código de curso passou a figurar nas planilhas de resultados do INEP, entretanto, para identificar um curso no período anterior, foi necessário criar um identificador composto de três elementos: código da IES, código da área e código do município.

Desde que passou a ser censitário em 2009, o Enade já foi realizado por mais de 3,3 milhões de estudantes (Tabela 2). Essa cifra equivale a aproximadamente 31% do total de alunos concluintes de graduação (10,8 milhões) em território nacional no período de 2009 a 2018. Levando-se em consideração que o ciclo de avaliação é trienal e que nem todos os cursos oferecidos são avaliados pelo Enade, essa cifra está dentro da esperada.

Tabela 1 – Número de estudantes concluintes inscritos e participantes no Enade a partir do ano de 2009

Ano	Número de estudantes concluintes inscritos	Número de estudantes concluintes participantes
2009	Não disponível	334.491
2010	160.737	149.590
2011	374.220	302.098
2012	587.351	469.460

Continua

Conclusão		
Ano	Número de estudantes concluintes inscritos	Número de estudantes concluintes participantes
2013	195.525	167.787
2014	481.720	396.862
2015	549.627	447.193
2016	216.044	195.423
2017	537.360	450.032
2018	550.845	462.242
Total	3.653.429	3.375.178

Fonte: Os autores (2022) adaptado de BRASIL (2021).

Comparando o número de participantes ao longo dos anos, percebe-se que o Ciclo I é aquele com o menor número de participantes, enquanto o Ciclo III registrou as maiores cifras, embora seja o Ciclo II aquele que apresente a maior quantidade e variedade de áreas avaliadas.

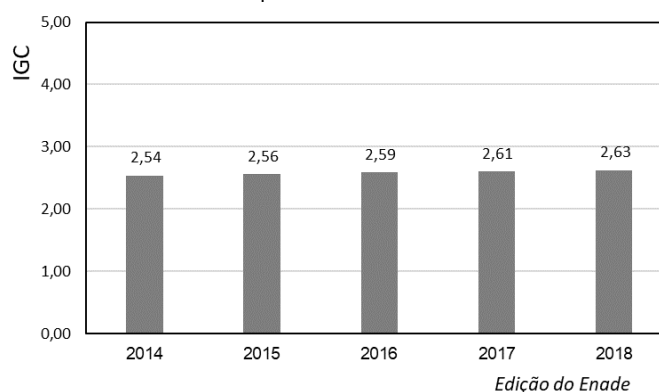
Resultados

Nesta seção, são apresentados os resultados da análise transversal múltipla, seguindo a descrição metodológica da seção anterior. Os resultados são apresentados separadamente para o IGC, CPC e seus componentes.

Análise do IGC

O IGC do conjunto de instituições considerado apresentou um leve e constante aumento durante as cinco edições ocorridas de 2014 a 2018, conforme mostra a Figura 2. O IGC contínuo médio oscila levemente à direita do centro do intervalo correspondente ao conceito 3, o que não deixa de ser esperado devido ao uso do afastamento padronizado.

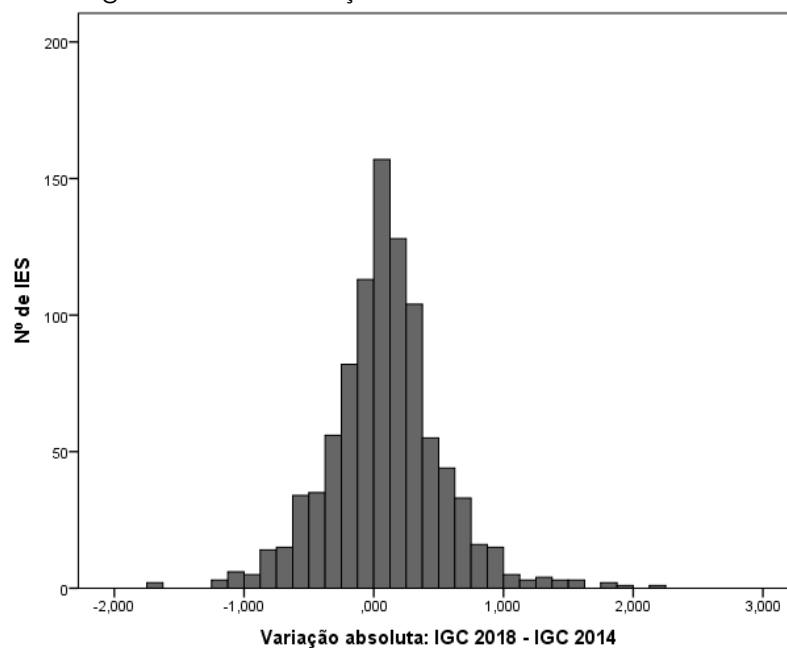
Figura 2 – Evolução do IGC no conjunto de 939 instituições consideradas nesta análise durante o período de 2014 a 2018



Fonte: Os autores (2022) adaptado de BRASIL (2021).

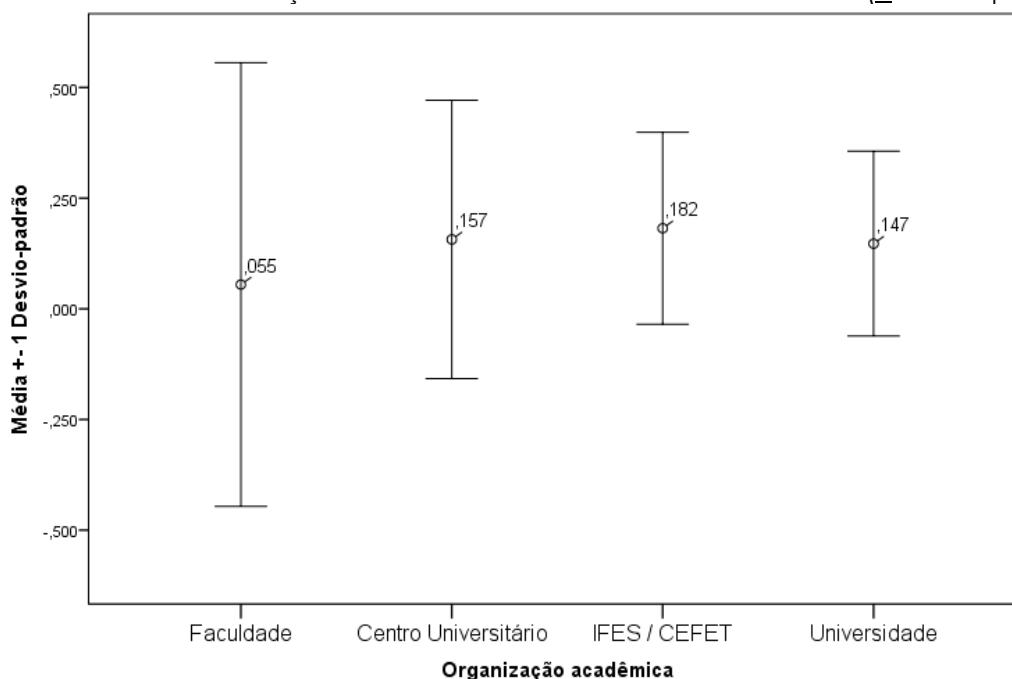
Das 939 instituições consideradas, 573 (61%) apresentaram incremento no IGC no período considerado, o que pode ser observado na Figura 3 com um aparente equilíbrio de diferenças positivas e negativas, mas que, na verdade, estão levemente concentradas à direita.

Figura 3 – Histograma das diferenças absolutas entre os IGCs de 2018 e 2014



Fonte: Os autores (2022) adaptado de BRASIL (2021).

A Figura 4 revela as diferenças na variação absoluta do IGC nas diferentes organizações acadêmicas. Percebe-se que nas faculdades, a diferença média foi a mais baixa, mas com um desvio-padrão superior. Já nos centros universitários, institutos e centros federais e nas universidades, as diferenças médias registradas correspondem a aproximadamente o triplo da cifra registrada nas faculdades. No geral, a diferença média ficou em torno de 0,09 pontos. O maior desvio-padrão registrado nas faculdades é esperado, visto que o número de cursos tende a ser menor neste tipo de instituição. Como o IGC é uma média entre os diferentes cursos, o que vemos é uma demonstração do funcionamento do Teorema Central do Limite (FISCHER, 2011), o qual diz que a variação da média é inversamente proporcional à raiz quadrada do número de observações.

Figura 4 – Médias das variações absolutas entre os IGCs de 2018 e 2014 (\pm desvio-padrão)

Fonte: Os autores (2022) adaptado de BRASIL (2021).

Em se tratando de um período sem mudanças metodológicas, pode-se afirmar que, em média, o indicador apresentou um leve aumento no período compreendido entre 2014 e 2018. Faz-se necessário investigar as causas por trás deste aumento.

A primeira suspeita foi a de que tenha havido um aumento relativo na presença da pós-graduação nas instituições de 2014 para 2018 e que este tenha sido o fator que impulsionou o aumento do IGC. Esta hipótese, contudo, foi refutada pois, independentemente da organização acadêmica, as variações no peso do componente da pós-graduação não explicam mais do que 3% das variações no IGC. Os componentes de mestrado e doutorado (M e D) explicam ainda menos: 0,3%. Portanto, as variações no IGC são quase que inteiramente explicadas pelas variações no componente G – a média dos conceitos preliminares de curso ponderada pelo número de matrículas.

Na próxima seção, a análise do CPC e seus insumos é apresentada.

Análise do Conceito Preliminar de Curso (CPC)

As mudanças nos pesos do CPC apresentadas no Quadro 1 tiveram impacto sobre os resultados do conceito. Entre 2007 e 2010, as mudanças nos pesos do CPC pouco afetaram os conceitos (BITTENCOURT; VIALI; RODRIGUES; CASARTELLI, 2010), mas, em contrapartida, as mudanças que passaram a vigorar em 2011 tiveram um

grande impacto sobre os resultados, o que pode ser identificado na Tabela 2. A substituição da nota dos ingressantes pelo resultado do Enem, a diminuição do peso do Enade e o aumento do IDD impulsionaram a distribuição dos conceitos à direita. Assim, constata-se um aumento no percentual de cursos classificados como 4 ou 5 a partir de 2011, o que permaneceu vigente depois da entrada do instrumento psicométrico em 2013 para quantificação dos insumos.

Tabela 2 – Distribuição dos Conceitos Preliminares de Curso (CPC) para todas as áreas avaliadas entre 2007 e 2018

Edição	CPC					% com CPC	Cursos
	1	2	3	4	5	4 ou 5	com CPC
2007	1,1%	23,9%	53,6%	19,0%	2,4%	21,3%	2.034
2008	1,1%	28,1%	50,2%	18,1%	2,5%	20,6%	4.820
2009	0,9%	33,0%	51,5%	13,2%	1,3%	14,6%	4.993
2010	0,7%	20,1%	53,9%	23,3%	2,0%	25,3%	2.762
2011	0,5%	14,0%	49,8%	32,5%	3,2%	35,6%	5.485
2012	0,3%	13,9%	57,4%	26,5%	1,9%	28,3%	5.351
2013	0,1%	10,4%	51,7%	35,2%	2,5%	37,7%	2.891
2014	0,3%	11,8%	56,7%	29,2%	2,0%	31,2%	7.479
2015	0,3%	10,6%	60,2%	27,6%	1,3%	28,9%	7.072
2016	0,4%	7,0%	50,5%	40,3%	1,9%	42,2%	4.197
2017	0,4%	9,1%	52,0%	36,3%	2,3%	38,5%	10.210
2018	0,4%	9,5%	56,6%	31,7%	1,7%	33,5%	8.520
Todos	0,5%	14,5%	54,1%	28,9%	2,0%	30,9%	65.814

Fonte: Os autores (2022) adaptado de BRASIL (2021).

Os resultados da Tabela 2 devem ser analisados com muito cuidado, pois são apenas uma organização geral de todos os CPCs atribuídos no período. Conforme já mencionado, há, sem dúvida, efeito de mudanças metodológicas. Também há mudanças no conjunto de cursos avaliados. Além disso, como cada componente do CPC é padronizado de acordo com a média e o desvio-padrão da área em cada edição, fica impossível avaliar a melhora ou piora do sistema através da distribuição dos conceitos em cada edição.

Dentre os atuais oito componentes do CPC, nem todos puderam ser comparados ao longo de todo o período de existência do indicador. Assim como no IGC, constatou-se que desde 2014 os critérios não sofreram alterações. Neste ponto, optamos por incluir também o ano de 2013, com a ressalva de que neste ano, o uso do ENEM não era individualizado. Deste modo, foram garantidas duas edições para

cada ciclo do Enade. A Tabela 3 mostra uma comparação dos resultados do CPC entre cursos que foram avaliados nas duas últimas edições do Enade para um mesmo ciclo.

Tabela 3 – Comparação dos resultados das faixas e dos valores contínuos do CPC considerando o mesmo conjunto de cursos nos pares de edições

Ciclo	Edições comparadas	Nº de cursos	Variação no CPC (faixa)			Kappa	R²
			Aumento	Igual	Diminuição		
I	2013 → 2016	2.890	26%	58%	17%	0,28	29%
II	2014 → 2017	7.022	25%	58%	17%	0,28	27%
III	2015 → 2018	6.108	21%	60%	19%	0,27	23%

Fonte: Os autores (2022) adaptado de BRASIL (2021).

A Tabela 3 permite inferir que aproximadamente 60% dos cursos tendem a repetir o CPC da edição anterior do Enade. O percentual de cursos com aumento de CPC tende a ser maior do que com diminuição, o que reforça os resultados gerais da Tabela 2, os quais apresentam uma tendência de aumento no CPC. Os valores do coeficiente Kappa indicam que há uma concordância regular (LANDIS; KOCH, 1977) entre os conceitos das duas edições e os valores do coeficiente de determinação (R²) indicam que de 23 a 29% das variações no CPC contínuo podem ser explicadas pelas variações na edição anterior.

Embora a padronização prejudique a análise do sistema como um todo por meio do CPC, há meios de contornar essa limitação analisando os escores brutos dos componentes, com exceção do Enade e do IDD. A comparação do Enade não é possível porque as provas são diferentes entre as edições. O IDD, como está relacionado ao Enade, também compartilha desta limitação. De qualquer forma, a comparação dos conceitos Enade e IDD entre duas edições sobre um mesmo conjunto de cursos não é vedada. A manutenção do conceito de uma edição para outra revela que, relativamente ao conjunto de instituições avaliado, o curso permaneceu dentro da mesma faixa. Isso quer dizer que um curso pode “ter corrido muito para permanecer no mesmo lugar”, lembrando a famosa frase de Lewis Carroll (LEWIS, 2017). Uma evolução do sistema em termos de desempenho na prova não tem como ser captada pela metodologia baseada no afastamento padronizado, mas a transição dos conceitos entre as edições pode ser feita (Tabelas 4 e 5).

Tabela 4 – Comparação dos resultados das faixas e dos valores contínuos do Enade considerando o mesmo conjunto de cursos nos pares de edições

Ciclo	Edições comparadas	Nº de cursos	Variação no Enade (faixa)			Kappa	R²
			Aumento	Igual	Diminuição		
I	2013 → 2016	2.890	28%	49%	23%	0,268	39%
II	2014 → 2017	7.022	28%	47%	25%	0,254	37%
III	2015 → 2018	6.108	25%	49%	26%	0,254	36%

Fonte: Os autores (2022) adaptado de BRASIL (2021).

De uma forma geral, pode-se inferir que aproximadamente metade dos cursos tendem a repetir o mesmo conceito Enade de uma edição para outra, enquanto aproximadamente a quarta parte registra aumento ou diminuição do indicador. Assim como no CPC, os valores do coeficiente Kappa indicam que há uma concordância regular entre os coeficientes das duas edições (LANDIS; KOCH, 1977). Os valores do coeficiente de determinação (R^2) indicam que um percentual considerável das variações do conceito Enade contínuo (de 36 a 39%) pode ser explicado pelas variações na edição anterior.

O conceito IDD, por sua vez, é o mais complexo dentre todos os componentes do CPC e também aquele que contribui com o maior peso. Embora Haar (2021) tenha apontado erros no cálculo do IDD desde 2014, decidiu-se pela manutenção da comparação do IDD neste estudo por dois motivos: 1) em todas as edições, desde 2014, o erro foi mantido; 2) o IDD calculado pelo INEP está diretamente correlacionado ao calculado pelo autor.

O cálculo do IDD envolve a realização de regressão onde o desempenho esperado para cada aluno é estimado a partir de um conjunto de variáveis preditoras. Assim, a diferença entre os resultados efetivamente observados e os estimados formam o indicador. O IDD, assim como o Enade, também sofre padronização. Detalhes técnicos sobre o cálculo do IDD podem ser obtidos em INEP (INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA 2018).

Tabela 5 – Comparação dos resultados das faixas e dos valores contínuos do IDD considerando o mesmo conjunto de cursos nos pares de edições

Ciclo	Edições comparadas	Nº de cursos	Variação no IDD (faixa)			Kappa	R²
			Aumento	Igual	Diminuição		
I	2013 → 2016	2.868	31%	39%	30%	0,087	9%
II	2014 → 2017	6.645	28%	46%	26%	0,134	5%
III	2015 → 2018	5.831	23%	51%	26%	0,163	5%

Fonte: Os autores (2022) adaptado de BRASIL (2021).

Os dados do IDD apresentam pequena concordância entre as edições, conforme revelam os coeficientes Kappa abaixo de 0,2. Como era esperado, a concordância mais baixa se deu entre 2013 e 2016 devido à diferença metodológica utilizada nos dois períodos, pois, em 2013, ainda não se utilizava a informação individualizada do Enem. Os coeficientes de determinação baixos indicam que é muito difícil prever o IDD de um curso a partir do conhecimento do resultado na edição anterior.

Os indicadores Enade e IDD são calculados a partir do desempenho na prova, mas há ainda dois conjuntos de indicadores que formam o CPC que permitem uma comparação entre as edições por meio da utilização de dados brutos. O primeiro deles refere-se a três indicadores derivados das 42 questões do questionário do estudante e que formam a dimensão "Percepção Discente sobre as Condições do Processo Formativo", cujas notas brutas variam de 1 a 6: Organização didático-pedagógica, Instalações e Infraestrutura e Oportunidades de Ampliação. As escalas de resposta para todas as questões do instrumento variam de 1 (discordância total) até 6 (concordância total), sendo esse último associado a melhor avaliação possível. A Tabela 6 apresenta a comparação das notas médias entre os pares de edição do Enade, considerando o mesmo conjunto de cursos. Esse pareamento permitiu a utilização do teste t de Student para amostras pareadas, o que aumenta a eficiência da comparação.

Tabela 6 – Comparação dos resultados das médias brutas dos três componentes da dimensão "Percepção Discente sobre as Condições do Processo Formativo"

Ciclo	Edições comparadas Ano <i>i</i> → <i>j</i>	n	Organização didático pedagógica			Infraestrutura			Oportunidades de Ampliação		
			Ano <i>i</i>	Ano <i>j</i>	Teste <i>t</i>	Ano <i>i</i>	Ano <i>j</i>	Teste <i>t</i>	Ano <i>i</i>	Ano <i>j</i>	Teste <i>t</i>
I	2013 → 2016	2.890	5,30	5,47	*	4,98	5,21	*	4,78	4,96	*
II	2014 → 2017	7.022	5,16	5,13	*	4,84	4,90	*	4,65	4,64	ns
III	2015 → 2018	6.108	5,35	5,18	*	5,20	5,15	*	4,89	4,77	*

* Diferença estatisticamente significativa ao nível de 5% pelo teste t de Student pareado bilateral.
Fonte: Os autores (2022) adaptado de BRASIL (2021).

As comparações da percepção discente revelam resultados diferentes entre os ciclos. O ciclo I, formado predominantemente por cursos da área da saúde, apresentou aumento nas médias de todos os três componentes. Já no ciclo II (engenharias, computação, licenciaturas, etc.) foram registradas as menores variações nas notas brutas dos três componentes. Organização didática teve uma

redução de 0,03 na média e Infraestrutura teve um aumento de 0,06 na média, ambas diferenças estatisticamente significativas. O item de Oportunidades de ampliação foi o único cujo teste t não resultou significativo. No ciclo III, todas os componentes apresentaram redução significativa nas médias.

Os resultados permitem inferir que, de um modo geral, a variação dos resultados de percepção discente apresentou comportamento distinto de um ciclo para outro. Destaca-se que as médias nacionais são muito elevadas em se tratando de uma escala ordinal de 6 pontos.

O segundo conjunto de indicadores se refere ao corpo docente. Os resultados de titulação e regime de trabalho dos professores são extraídos pelo INEP diretamente do Censo da Educação Superior. Os resultados brutos permitem comparar o percentual de professores de acordo com a titulação e regime de trabalho para uma avaliação temporal sem o efeito de padronizações (Tabela 7).

Tabela 7 – Comparação dos resultados das proporções médias de titulação e regime de trabalho docente entre as edições

Ciclo	Edições comparadas Ano <i>i</i> → Ano <i>j</i>	n	Mestrado		Doutorado		Regime de trabalho	
			Ano <i>i</i>	Ano <i>j</i>	Ano <i>i</i>	Ano <i>j</i>	Ano <i>i</i>	Ano <i>j</i>
I	2013 → 2016	2.868	81,0% ^a	86,6% ^b	39,0% ^a	49,8% ^b	81,3% ^a	85,8% ^b
II	2014 → 2017	6.645	83,8% ^a	88,8% ^b	45,7% ^a	53,9% ^b	84,3% ^a	86,8% ^b
III	2015 → 2018	5.831	86,7% ^a	90,2% ^b	43,3% ^a	46,5% ^b	84,8% ^a	86,7% ^b

Obs.: Proporções médias seguidas de mesma letra dentro de um mesmo componente em um mesmo ciclo indicam que não existe diferença significativa ao nível de 5%.

Fonte: Os autores (2022) adaptado de BRASIL (2021).

Os resultados revelam que a titulação docente e o percentual de professores em regime de tempo integral ou parcial aumentaram em todos os três ciclos. Os aumentos foram mais expressivos nos ciclos I e II.

No ciclo I, por exemplo, o aumento do percentual médio de professores com doutorado foi superior a 10 pontos percentuais. Esses resultados não são surpreendentes, porque a pós-graduação brasileira *stricto sensu* vem registrando uma série de titulações crescente. Nos últimos 20 anos, foram mais de 1 milhão de mestres e doutores (CABRAL; SILVA; PACHECO; MELO, 2020), dos quais muitos foram absorvidos por instituições acadêmicas.

Considerações finais

Este estudo teve por objetivo apresentar uma análise comparativa dos resultados do Enade desde a divulgação dos primeiros indicadores. Devido às mudanças metodológicas ocorridas ao longo do período e o efeito da padronização, houve limitação em relação aos períodos que puderam ser comparados.

Os resultados permitiram constatar que, entre 2014 e 2018, o IGC apresentou um pequeno acréscimo médio de 0,09 pontos causados principalmente pelo aumento nos CPCs dos cursos de graduação. Dentre os componentes do CPC, o Enade e o IDD tendem a estabilidade, como esperado, enquanto àqueles associados ao corpo docente experimentaram aumento durante o período de implantação do SINAES.

Diante desses resultados, surge um importante questionamento: por que alguns indicadores experimentaram aumento, se todos passam por um mesmo processo de padronização? A resposta está na distribuição truncada. Quando as notas médias são muito elevadas, significa que muitas instituições atingiram o valor máximo possível (5), ocorrendo o truncamento. Isso ocorreu com o percentual de professores com mestrado e o percentual de professores em regime de tempo de trabalho integral ou parcial. Assim, a média dos escores padronizados ficou bem acima do centro da escala, que seria 2,5. Essa aproximação do máximo da escala pode estar indicando que, no futuro, esses indicadores de titulação docente deixem de ser necessários, o que já é uma realidade em muitas avaliações internacionais.

O efeito de truncamento também ocorre nos indicadores de percepção discente, ainda que em menor grau. À primeira vista, encontrar médias altas nesses indicadores parece ser um aspecto positivo, mas, do ponto de vista do poder discriminativo das escalas, temos um problema porque a variância é muito baixa.

Uma análise detalhada dos microdados do Enade revela que algumas instituições obtiveram nota máxima em todas as questões do instrumento por parte de todos os alunos participantes do Enade, o que pode ser considerado um tanto improvável, apesar de possível. Esta constatação pode ser indício de alguma espécie de artificialização dos resultados.

Conclui-se que os resultados do Enade não permitem uma avaliação temporal, o que o torna um instrumento inadequado para avaliar evolução ou involução do sistema ao longo do tempo. Contudo, individualmente, cada instituição pode ter uma boa noção do desempenho de seus cursos em cada edição

transversalmente. O estudo revelou, também, que aproximadamente 40% dos cursos deverão ter CPCs diferentes da edição anterior. O IDD, principal componente do CPC, mostrou-se o mais difícil de prever a partir de resultados anteriores.

Por fim, este estudo denota a importância da revisão deste processo de avaliação externa, o que já foi desencadeado pela Diretoria de Avaliação da Educação Superior do INEP em 2019 (BRASIL, 2019) mas, no entanto, foi interrompido pela alternância na gestão institucional.

Agradecimentos

Este trabalho teve apoio financeiro do CNPq, Chamada Universal 422215/2016-3.

Referências

BARREYRO, G. B. De exames, rankings e mídia. *Avaliação: Revista de Avaliação do Ensino Superior*, Campinas, SP, v. 13, n. 3, p. 863-868, 2008. DOI: <https://doi.org/10.1590/S1414-40772008000300017>. Disponível em: <https://www.scielo.br/j/aval/a/qXYXfm3tM5YrVFDYQDxXhN/abstract/?lang=pt#>. Acesso em: 7 maio 2021.

BARREYRO, G. B.; ROTHEN, J. C. Percurso da avaliação da educação superior nos Governos Lula. *Educação e Pesquisa*, São Paulo, v. 40, n. 1, p. 61-76, jan./mar. 2014. DOI: <https://doi.org/10.1590/S1517-97022014000100005>. Disponível em: <https://www.scielo.br/j/ep/a/fLsPHntHb77vRNtNpdLG5Jv/>. Acesso em: 7 maio 2021.

BITTENCOURT, H. R., CASARTELLI, A. O.; RODRIGUES, A. C. M. Sobre o índice geral de cursos (IGC). *Avaliação: Revista de Avaliação do Ensino Superior*, Campinas, SP, v. 14, n. 3, p. 667-682, 2010.

BITTENCOURT, H. R.; VIALI, L.; CASARTELLI, A. O.; RODRIGUES, A. C. M. Uma análise da relação entre os conceitos Enade e IDD. *Estudos em Avaliação Educacional*, São Paulo, SP, v. 19, n. 40, p. 247-262, maio/ago. 2008.

BITTENCOURT, H. R.; VIALI, L.; RODRIGUES, A. C. M.; CASARTELLI, A. O. Mudanças nos pesos do CPC e seu impacto nos resultados de avaliação em universidades federais e privadas. *Avaliação: Revista de Avaliação do Ensino Superior*, Campinas, SP, v. 15, n. 3 p. 147-166, 2010. DOI: <https://doi.org/10.1590/S1414-40772010000300008>. Disponível em: <https://www.scielo.br/j/aval/a/SJTW3vrXqCpMpFCX63vXbfS/?lang=pt#>. Acesso em: 7 maio 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Sistema nacional de avaliação da educação superior (SINAES)*. v. 1. Brasília, DF: Inep, 2011. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_superior/sinaes_2011_volume_1_analise_dos_resultados_do_enade_ciencias_saude_ciencias_agrarias_servivo_social_2004_e_2007.pdf. Acesso em: 7 maio 2021.

BRASIL. Lei n.º 10.861, de 14 de abril de 2004. Institui o Sistema Nacional de Avaliação da Educação Superior - SINAES e dá outras Providências. *Diário Oficial da União*, Brasília, DF, 15 abr. 2004.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Censo da educação superior 2021: notas estatísticas*. Brasília, DF: Inep, 2022. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_da_educacao_superior_2021.pdf. Acesso em: 7 maio 2022.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. *Indicadores de qualidade da educação superior*. Brasília, DF, 2021. Disponível em: <https://www.gov.br/inep/pt-br/acesso-a->

informacao/dados-abertos/indicadores-educacionais/indicadores-de-qualidade-da-educacao-superior. Acesso em: 23 fev. 2021.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Portaria INEP nº 748, de 23 de agosto de 2019. Institui a Comissão de Assessoramento para Revisão dos Processos Avaliativos, Instrumentos e Técnicas de Coleta de Dados e Indicadores da Educação Superior do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Diário Oficial da União*, BRASÍLIA, DF, 30 ago. 2019.

BRASIL. Ministério da Educação. Portaria Normativa n. 40, de 29 de dezembro de 2010. 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 União*, Brasília, DF, 29 dez. 2010.

CABRAL, T. L. de O.; SILVA, F. C. da; PACHECO, A. S. V.; MELO, P. A. de. A CAPES E suas sete décadas: trajetória da pós-graduação stricto sensu no Brasil. *Revista Brasileira de Pós-Graduação*, [S. l.], v. 16, n. 36, p. 1–22, 2020. DOI: <https://doi.org/10.21713/rbpg.v16i36.1680>. Disponível em: <https://rbpg.capes.gov.br/rbpg/article/view/1680>. Acesso em: 15 set. 2023.

CORTELAZZO, A. L.; RIBEIRO, V. K. Enade 2005 e 2008: desempenho dos estudantes de biologia de instituições de educação superior estaduais e municipais de São Paulo. *Ciência & Educação*, Bauru, v. 19, n. 2, p. 409-424, 2013. DOI: <https://doi.org/10.1590/S1516-73132013000200012>. Disponível em: <https://www.scielo.br/j/ciedu/a/j59N7hYYmsjtDqBgDSS9WMx/abstract/?lang=pt>. Acesso em: 7 maio 2021.

FISCHER, H. *A history of the central limit theorem: from classical to modern probability theory*. [S. l.]: Springer: 2011.

GRIBOSKI, C. M. O Enade como indutor da qualidade da educação superior. *Estudos em Avaliação Educacional*, São Paulo, v. 23, n. 53, p. 178-195, 2012. DOI: <https://doi.org/10.18222/eae235320121920>. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/1920>. Acesso em: 7 maio 2021.

GURGEL, C. R. Análise do exame nacional de desempenho dos estudantes após o primeiro ciclo avaliativo das áreas de agrárias, saúde e serviço social do estado do Piauí. *Ensaio: Aval. Pol. Públ. Educ.*, Rio de Janeiro, v. 18, n. 66, p. 85-104, 2010. DOI: <https://doi.org/10.1590/S0104-40362010000100006>. Disponível em: <https://www.scielo.br/j/ensaio/a/5Y4TqSGX767DfpZTbwBCXzD/abstract/?lang=pt#>. Acesso em: 7 maio 2021.

HAAR, E. Um método correto para calcular o indicador de valor agregado IDD do sistema nacional de avaliação da educação superior. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, SP, v. 26, n. 2, p. 568-586, 2021. DOI: <https://doi.org/10.1590/S1414-40772021000200012>. Disponível em:

<https://www.scielo.br/j/aval/a/SzvcwJ7wnJLjCNXXz4d8pXv/abstract/?lang=pt#>. Acesso em: 7 ago. 2021.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. *Nota Técnica nº 17/2018/CGCQES/DAES*. Brasília, DF: Inep, 2018. Disponível em: https://download.inep.gov.br/educacao_superior/enade/notas_tecnicas/2017/nota_tecnica_n17_2018_calculo_idd.pdf. Acesso em: 7 maio 2021.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. *Anais dos seminários regionais sobre autoavaliação institucional e comissões próprias de avaliação (CPA)*. Brasília: INEP, 2015. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_superior/anais_dos_seminarios_regionais_sobre_autoavaliacao_institucional_e_comissoes_proprias_da_avaliacao_cpa_2013.pdf. Acesso em: 7 maio 2021.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. *Biometrics*, Washington, v. 33, n. 1, p. 159-174, 1977. DOI: <https://doi.org/10.2307/2529310>. Disponível em: <https://www.jstor.org/stable/2529310>. Acesso em: 7 ago. 2021.

LEWIS Carol 1832-1898. In: RATCLIFFE, S. (ed.). *Oxford essential quotations*. Oxônia, Reino Unido: Oxford University Press, 2017. Disponível em: <https://www.oxfordreference.com/display/10.1093/acref/9780191843730.001.0001/q-oro-ed5-00002622?rsk=yw5VTNV&result=646>. Acesso em: 7 maio 2021.

MARCHELLI, P. S. O sistema de avaliação externa dos padrões de qualidade da educação superior no Brasil: considerações sobre os indicadores. *Ensaio: Aval. Pol. Públ. Educ.*, Rio de Janeiro, v. 15, n. 56, p. 351-372, 2007.

MCCOWAN, T.; BERTOLIN, J. *Inequalities in higher education access and completion in Brazil*. Geneva: UNRISD, 2020.

PEREIRA, C. A; MACHADO-TAYLOR, M. L.; ARAUJO, J. F. E. F.; TEIXEIRA, A. F. Análise longitudinal do desempenho dos estudantes da educação superior brasileira: accountability pelo desempenho dos egressos. *Revista FORGES*, Iléus, v. 5, n. .2, p. 71-98, 2017.

POLIDORI, M. M. Políticas de avaliação da educação superior brasileira: Provão, SINAES, IDD, CPC, IGC e... outros índices. *Avaliação: Revista da Avaliação da Educação Superior*, Campinas, v. 14, n. 2, p. 439-452, 2009. DOI: <https://doi.org/10.1590/S1414-40772009000200009>. Disponível em: <https://www.scielo.br/j/aval/a/yFb9SmwXsdtq9rrzTp3fhFs/abstract/?lang=pt>. Acesso em: 7 maio. 2021.

RISTOFF, D.; GIOLO, J. O Sinaes como sistema. *Revista Brasileira de Pós-Graduação*, [S. l.], v. 3, n. 6, 2011. DOI: 10.21713/2358-2332.2006.v3.106. Disponível em: <https://rbpg.capes.gov.br/rbpg/article/view/106>. Acesso em: 7 maio 2021.

VERHINE, R. E. Avaliação e regulação da educação superior: uma análise a partir dos primeiros 10 anos do SINAES. *Avaliação: Revista da Avaliação da Educação*

Superior, Campinas, v. 20, n. 3, p. 603-619, 2015. DOI: <https://doi.org/10.1590/S1414-40772015000300003>. Disponível em: <https://www.scielo.br/j/aval/a/n8WYbvtmRRgBFtvr3QkcKCt/?format=pdf&lang=pt>. Acesso em: 7 ago. 2021.