

Importância da leitura e gerenciamento do Banco Nacional de Itens na elaboração de avaliação institucional: desafios e dificuldades

The importance of the National Item Bank in the elaboration of institutional evaluation: challenges and difficulties

Importancia del Banco Nacional de Elementos en el desarrollo de la evaluación institucional: retos y dificultades

Atair Silva de Sousa

Universidade de São Paulo, Faculdade de Educação, São Paulo, São Paulo, Brasil

atair.sousa@usp.br | <https://orcid.org/0000-0002-1355-8797>

Resumo

Este artigo discute as práticas pedagógicas envolvidas na produção de itens de qualidade para bancos de itens de avaliação institucional. O objetivo central consistiu em verificar os desafios e as dualidades para produção de item de qualidade para seu efetivo aproveitamento e utilização em avaliações institucionais ou para compor seus respectivos bancos de itens. Foram realizadas uma revisão bibliográfica documental e uma busca de dados e informações disponibilizadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep). Além disso, foi necessário realizar uma revisão da literatura pertinente ao tema para identificar conceitos e definição acerca de avaliação em larga escala e/ou institucional. Foram coletadas respostas de questionário disponibilizado via Google forms para docentes elaboradores e revisores de itens. Os resultados apontaram para falhas nos processos e falta de experiência de parte dos docentes na elaboração de itens para avaliação em larga escala.

Palavras-chave: Avaliação em larga escala. Banco de itens. Políticas educacionais. SINAES.

Abstract

This paper discuss the pedagogical practices involved in the production of quality items for institutional assessment item banks. The central aim was to verify the challenges and dualities in producing quality items for their effective use in institutional assessments or to compose their respective item banks. A documental bibliographic review and a search for data and information made available by the National Institute for Educational Studies and Research (INEP) were conducted. In addition, it was necessary to conduct a literature review pertinent to the topic to identify concepts and definition about large-scale and/or institutional evaluation. Responses were collected from a questionnaire made available via Google forms to faculty members who were item designers and reviewers. The results pointed to flaws in the processes and lack of experience of part of the teachers in the preparation of items for large-scale assessment.

Keywords: Large-scale assessment. Item bank. Educational policies. SINAES.

Resumen

En este artículo es analizan las prácticas pedagógicas que intervienen en la producción de ítems de calidad para los bancos de ítems de evaluación a institucional. El objetivo principal era verificar los retos y dualidades en la producción de ítems de calidad para su uso efectivo en las evaluaciones institucionales o para componer sus respectivos bancos de ítems. Se realizó una revisión bibliográfica documental y una búsqueda de datos e información disponible en el Instituto Nacional de Estudios e Investigaciones Educativas (INEP). Además, fue necesario realizar una revisión de la literatura pertinente al tema para identificar los conceptos y la definición

Artigo recebido em: 16/09/2023 | Aprovado em: 03/08/2023 | Publicado em: 15/11/2023

Como citar:

SOUSA, Atair Silva de. Importância da leitura e gerenciamento do Banco Nacional de Itens na elaboração de avaliação institucional: desafios e dificuldades. **Pesquisa e Debate em Educação**, Juiz de Fora: UFJF, v. 13, p. 1-16, e35310. 2023. ISSN 2237-9444. DOI: <https://doi.org/10.34019/2237-9444.2023.v13.35310>.

sobre la evaluación a gran escala y/o institucional. Las respuestas se recogieron a partir de un cuestionario puesto a disposición a través de formularios de Google para los profesores que preparan y revisan los artículos. Los resultados apuntaron a fallos en los procesos y a la falta de experiencia de una parte de los profesores en la preparación de ítems para la evaluación a gran escala.

Palabras clave: Evaluación a gran escala. Banco de artículos. Políticas educativas. SINAES

1 Introdução

A partir da Lei de Diretrizes e Bases da Educação (LDB/96), o processo de avaliação da qualidade da educação se consolidou dentro da premissa do dever do Estado com a avaliação educacional (BRASIL, 1996). Nesse sentido, delegando a União de assegurar o processo avaliativo da trajetória acadêmica. No que tange à educação superior. Lima et al. (2020, p. 263) propõem que, além disso, “objetivando a definição de prioridades e a melhoria da qualidade da educação, assegurar processo nacional de avaliação das Instituições de Educação Superior (IES); e autorizar, reconhecer, credenciar, supervisionar e avaliar os cursos dessas IES”.

No contexto da LDB/96 as transformações estruturaram a educação nacional, fato que deu relevo aos processos de avaliação com vistas à melhoria da qualidade da educação. Isso proporcionou recurso para a regulação do setor, a acreditação de instituições da Educação Superior (IES) e de cursos. Na educação superior, por exemplo, a avaliação da qualidade assumiu lugar especial dentro das políticas educacionais. O Ministério da Educação, como órgão institucional, adotou ações concretas na orientação de diretrizes educacionais.

O tipo de instrumento avaliativo necessário para atender as demandas desse processo de avaliação da qualidade da educação, consiste de avaliação em larga escala, já que as reformas educacionais implementadas nas décadas de 80, 90 e as duas primeiras do século XXI demandaram a utilização de instrumento que avaliasse a gestão de redes de ensino.

Tendo que, inicialmente, a inserção da avaliação em larga escala no contexto da escola, objetivou fortalecer os processos educativos, sem vincular o resultado às atividades do professor em sala, não o responsabilizando pelos resultados obtidos pelos alunos. Entretanto, estabelece um modelo de avaliação que leva em conta o princípio da *accountability*, tendo em vista a mensuração de formas possíveis de modo a organizar a escola, no que se refere à docência, à aprendizagem e da organização curricular (WELTER; PASINI, 2014).

Somente a partir da década de 1990, segundo Amaro (2016), que “a avaliação em larga escala, no Brasil, intensifica-se e se impõe como motor na engrenagem educacional sob o pretexto de promoção da qualidade da educação em seus diversos níveis” (AMARO, 2016, p. 464). Tendo em vista que, “a possibilidade de essa avaliação atender mais claramente às necessidades políticas de reforma tornou-a prioridade e modelo que passou a orientar a avaliação das aprendizagens, avaliação da escola e de currículo” (SOUSA; FERREIRA, 2019, p. 15). Com isso, se apresentando como um novo elemento que foi incorporado à avaliação educacional: as avaliações externas.

De acordo com Bauer, Alavarse e Oliveira (2021) “há ainda autores que discutem que algumas implicações das avaliações sobre as redes e escolas não são,

em essência, boas ou ruins, mas que podem ter influências diferenciadas, a depender de cada contexto escolar” (BAUER; ALAVARSE; OLIVEIRA, 2021, p. 1377).

As avaliações organizadas e gerenciadas por instituição externa à escola são denominadas externas. Diferentes daquelas que são conduzidas pelos próprios professores em exercício na escola, que são denominadas internas. As avaliações externas podem ser utilizadas em larga escala, quando utilizadas para avaliar sistemas de ensino, por exemplo. Neste caso, fica evidente sua visibilidade em decorrência de utilidade em política pública educacional (ALAVARSE, 2013).

A avaliação externa em larga escala possui relevante utilidade no contexto da regulação dos sistemas de ensino. Suas metodologias são apropriadas na geração de indicadores para a gestão do ensino e da escola. Este tipo de avaliação possui um caráter sistêmico e, por conseguinte, assume o papel de avaliação sistêmica. Isso ocorre em função de seu poder de avaliar uma rede ou sistema de ensino. Neste caso é considerada em larga escala por envolver um grande número de estudantes no processo avaliativo.

Conforme assinalado, a avaliação em larga escala é aplicada para um quantitativo razoável de estudantes que vai para além do âmbito de uma turma, curso ou mesmo de uma instituição, ou seja, sua aplicação somente se justifica para redes ou sistemas. Ela possui padrão específico com propósitos em três frentes: relacional (objeto), caracterização (instrumental) e mensuração (resultados).

Para o MEC essas avaliações, destinadas ao ensino básico, contribuem para a melhoria da qualidade da educação e para o combate das desigualdades, além da democratização da educação pública (SOUSA; ARCAS, 2010). Ademais, “[...] espera-se que induzam ao desenvolvimento de uma cultura avaliativa que estimule o controle social sobre os processos e resultados de ensino” (SOUSA; ARCAS, 2010, p. 183).

Além disso, espera-se que a avaliação externa em larga escala possa contribuir com a qualidade da educação, pois seria por meio dela que se avalia os processos educacionais, desde o desempenho dos estudantes inserido na trajetória de sua formação, passando pelos testes de suas proficiências, até os diagnósticos das redes e sistemas de ensino (WIEBUSCH, 2012).

Welter e Werle alertam para o fato de que, ao analisar o discurso “[...] do político pedagógico sobre as avaliações de larga escala, destaca-se que a sociedade brasileira foi desafiada a construir a base nacional curricular comum (BNCC), com foco na elaboração de objetivos de aprendizagem, colocada à consulta pública em 2015” (WELTER; WERLE, 2021, p. 445). Para o MEC as avaliações em larga escala, destinadas a educação superior têm propósitos semelhantes, embora, mais abrangentes.

Diante do exposto e postas todas essas características, conceitos e particularidades, este estudo busca responder as seguintes questões: quais desafios são postos na produção de itens para composição das avaliações em larga escala? Que tipos de dificuldades são encontrados na elaboração de itens de qualidade para essas avaliações?

O objetivo central consistiu em verificar os desafios e as dualidades para produção de item de qualidade para seu efetivo aproveitamento e utilização em avaliações institucional ou para compor seus respectivos bancos de itens.

Para apresentação deste estudo, este artigo ficou estruturado em 5 (cinco) seções além desta Introdução. Na seção 2 (dois) são apontados os procedimentos metodológicos. Na seção 3 (três) é apresentado o Banco Nacional de Itens (BNI) como instrumento condicional na elaboração de avaliações em largas escalas e necessário para processo de qualidade na instrumentação da avaliação da educação. Na seção 4 (quatro) objetivou-se dar ênfase à análise de resultados e propostas pedagógicas que auxiliem na produção de item de qualidade para avaliação em larga escala. Na seção 5 (cinco) estão as considerações finais apresentando uma síntese da pesquisa.

2 Formatação

Esta pesquisa consistiu de uma metodologia de natureza propositiva, com uma abordagem mista e procedimentos bibliográfico e documental. O método qualitativo predominou diante das análises, diálogos e discussões das assertivas contextualizadas aos aspectos pedagógicos da produção de itens para avaliação em larga escala. Os dados coletados são tantos primários quanto secundários a depender de suas fontes, porém ambos de natureza quantitativa, valores inteiros relativos aos respondentes de questionário e número de itens do BNI do Instituto Nacional de Estudos e Pesquisas Educacionais – Anísio Teixeira – (BNI-INEP).

A fonte de dados primários utilizados neste estudo teve sua origem em um formulário com cinco questões respondidas por docentes do banco de avaliadores de itens do Inep. Para este estudo, optou-se por um conjunto restrito de apenas 20 docentes respondentes, em função de características específicas deste estudo, tais como: ser um estudo preliminar e servir de pré-testagem; servir de estudo base para estudos de maior abrangência; e restrição de tempo. Desse conjunto de docentes, 15 responderam ao questionário no prazo e de forma integral. Para esse conjunto de docentes, não houve qualquer restrição em relação a qual exame o docente participa ou participou como elaborador/revisor de itens, já que o Inep possui variadas bases de elaboradores de itens.

Já os dados secundários foram disponibilizados pela Ouvidoria do Inep, cuja solicitação teve como número de protocolo: 23546.025418/2021-09 (BRASIL, 2021). As análises estatísticas demandaram o uso de ferramentas descritivas simples diante da coleta de dados, sejam na construção de tabelas, quadros e gráficos, dados percentuais ou processos de contagem.

Com esse entendimento, tomou-se como objeto a análise de dados secundários e informações disponibilizadas entre 2010 e 2019, com a hipótese de mapear as ações previstas no processo de capacitação dos colaboradores do BNI. O período entre 2010 e 2019 se justifica pelo período de vigência do BNI e dos últimos resultados divulgados pelo Inep até a finalização deste estudo. O BNI foi implantado pelo Inep em 2010 e análise e discussão deste estudo se restringiu aos últimos resultados divulgados, o que ocorreu em 2019. Além disso, foi necessário realizar uma revisão da literatura pertinente ao tema. Observando-se as especificidades da base, optou-se pela utilização dos dados do Inep, em função desse órgão apresentar coletas de dados de todo o processo e atividades do BNI.

Por meio dos dados coletados foram construídos gráficos e tabelas cujo valores foram transformados em valores inteiros, tanto nos valores referentes aos itens quanto às percentagens relativas às respostas do questionário. Entende-se que esse procedimento poderia facilitar futuras comparações entre o período avaliado e as que o sucederem. Neste caso, visualizar a evolução da qualidade de itens aproveitados, se de fato, evoluíram em quantidade e qualidade.

3 O BNI como processo de qualidade na instrumentação da avaliação da educação

Ao se referir ao BNI, tem-se que:

O BNI é essencial para a elaboração e a aplicação das avaliações nacionais em larga escala da educação, que se distinguem das avaliações internas por não serem feitas pelo professor ou pela própria instituição de ensino para avaliar e propor alternativas no âmbito da sala de aula ou da instituição. Elaboradas por um órgão externo às escolas, como o Inep, as avaliações em larga escala têm a finalidade de fazer juízos de valor e propor alternativas em âmbito mais amplo que o da instituição de ensino. Além das avaliações em larga escala, o BNI também atende à demanda por itens para os exames do Inep. Atualmente, o BNI do Inep reúne um acervo de itens para a realização de variados instrumentos avaliativos. No Quadro 1 são apresentados os instrumentos avaliativos que utilizam o BNI (INEP, 2020, s.p).

Quadro 1: Processos dos instrumentos avaliativos do BNI

PROCESSO DO BNI	
Instrumento Avaliativo	Nível Educacional
Provinha Brasil; Sistema Nacional de Avaliação da Educação Básica - SAEB.	Educação Básica
Programa Internacional de Avaliação de Alunos – Pisa.	Educação Básica
Exame Nacional do Ensino Médio – Enem.	Educação Básica
Exame Nacional de Desempenho de Estudantes – Enade.	Educação Superior
Exame Nacional de Certificação de Competências de Jovens e Adultos – ENCCEJA.	Educação Básica
Revalidação dos Diplomas Médicos – REVALIDA.	Educação Superior
Certificado de Proficiência na Língua Brasileira de Sinais – PROLIBRAS.	Educação Básica
Certificado de Proficiência em Língua Portuguesa - CELPE-BRAS.	Educação Básica

Fonte: Sousa (2021b, p.6).

De acordo como o órgão responsável pelo BNI, o Inep,

É a manutenção de um banco de itens que oferece subsídios para construção dos testes e questionários. Com o BNI, o Inep garante uma quantidade expressiva de itens com comprovada qualidade técnico-pedagógica e psicométrica para compor testes de avaliações em larga escala e exames. O BNI define-se, portanto, como uma coleção de itens de testes de natureza específica – organizados segundo determinados critérios – disponíveis para a construção de instrumentos de avaliação (INEP, 2021a, s.p).

Nesse contexto da manutenção do BNI, a entrada constante de itens de qualidade seria fundamental para sua existência. Por isso a necessidade da colaboração de professores e pesquisadores da educação brasileira são chamados para a construção itens que possam fazer parte desse banco. Neste caso, o Inep realiza chamadas públicas e seleciona colaboradores interessados em elaborar e revisar itens para o BNI. Esse procedimento formata o Banco de Colaboradores do BNI para cada avaliação ou exame a ser desenvolvido no âmbito do Inep.

No caso da educação superior, o seu banco de itens,

[...] foi criado em 2010 e tem por objetivo reunir questões de qualidade técnica que permitam ao Inep avaliar por meio do exame o conhecimento dos estudantes em relação aos conteúdos programáticos, habilidades e competências previstas nas Diretrizes Curriculares Nacionais (DCNs) dos cursos de graduação. O Inep busca qualificar o exame com a construção de itens elaborados por profissionais experientes. O BNI é um sistema importante na construção dos instrumentos avaliativos do Inep. Isso porque fornece a segurança necessária ao processo de construção das provas, confiabilidade e garante maior participação da comunidade acadêmica na construção dos Itens (INEP, 2021a, s.p).

Portanto, o BNI da educação superior possui o formato de um sistema computacional no qual permanecem armazenados os itens de testes de natureza específica que podem ser aproveitados em exames realizados pelo Inep. Esses itens são organizados seguindo critérios estipulados por cada edição de cada exame (CAMPOS, 2013). Esse banco consiste de “um acervo de itens elaborados a partir de matrizes de conteúdo, competências e habilidades, pré-definidos, buscam estimar com a maior precisão possível a proficiência dos estudantes com relação aos conteúdos de seus respectivos cursos de graduação” (INEP, 2021a, s.p).

4 Discussões e análise de resultados

Com a implantação da avaliação institucional na educação superior brasileira, de acordo com Cerri (2016), foi constatada uma maior participação da comunidade acadêmica na construção dos instrumentos avaliativos, respeitando os princípios estabelecidos pelas normativas educacionais. Essa maior participação dos acadêmicos, tornou-se possível a construção de um processo continuado de no fluxo de processos com inclusão de novas etapas, tais como: revisão final, realizada pelos docentes representantes de cada área de conhecimento, respectivamente, e

de membros das Comissões Assessoras de Áreas (CAAs) no caso de cursos de graduação, tenho em vista a revisão de sensibilidade de cada item (CERRI, 2016).

A prova do Enade é caracterizada como uma avaliação em larga escala. Com isso, considera-se válido tomá-la como um exemplo para avaliar a produção de itens para esse tipo de avaliação. Considerando o BNI-Enade, na Tabela 1 está apresentado o quantitativo de itens elaborados, aceitos e descartados por modalidade (bacharelado, Tecnológico e Formação Geral), na edição do Enade 2019 (SOUSA, 2021b).

Tabela 1: Índice de aproveitamento do BNI-Enade 2019

APROVEITAMENTO DO BNI-ENADE EM 2019				
Modalidade	Elaborados	Aceitos	Descartes	Aproveitamento (%)
Bacharelado	3.542	2.310	1.232	65
Tecnológico	883	480	403	54
Formação Geral	58	29	29	50
Total	4.483	2.819	1.664	63

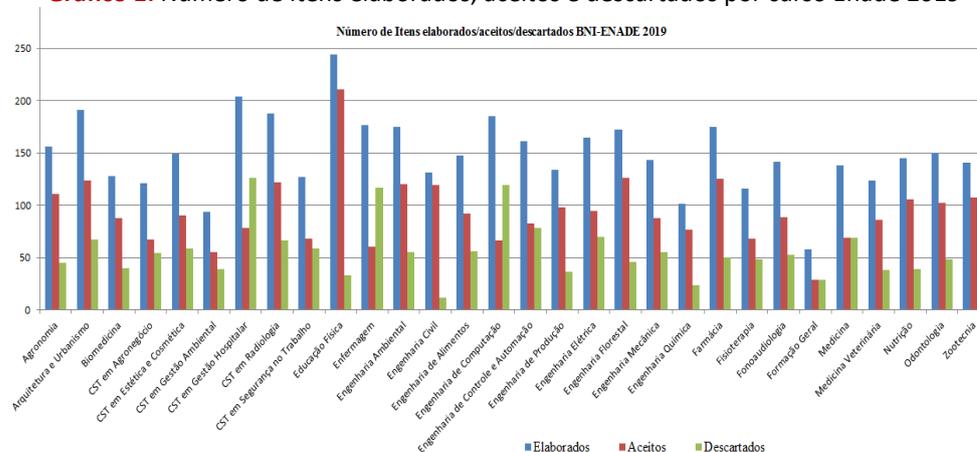
Fonte: Sousa (2021b, p. 15).

O resultado do aproveitamento total de itens produzidos (4.483) aponta para um número considerado, satisfatório, de itens aprovados (2.819). Por outro lado, leva a um percentual de 37% de itens reprovados. O número de itens aprovados pode ser considerado satisfatório ao levar em consideração de que, “para cada prova, são utilizados 40 itens, sendo 30 itens de Conhecimentos Específicos e 10 itens de Formação Geral” (INEP, 2021b, s.p). Contudo, “é necessário estar atento ao abastecimento do BNI-Enade e desenvolver propostas contínuas de aprimoramento nos processos de capacitação para que os resultados não sofram declínio” (SOUSA, 2021b, p. 15).

O gráfico a seguir, apresenta a distribuição dos itens segundo as áreas de conhecimento. Percebe-se que o processo de inserção de itens no BNI-Enade ocorre de forma contínua e uniforme, de tal modo que, o curso de capacitação disponibilizado aos colaboradores possui seu formato em metodologias ativas e propositivas, já que os resultados advindos desse curso são promissores. Chama-se a atenção para o fato de que esse curso e suas metodologias não se configuram como um processo pronto e acabado. Com isso, ratifica-se a necessidade de manutenção e continuidade dos processos de melhorias na formação cada vez mais atualizadas dos colaboradores (SOUSA, 2021b).

No gráfico a seguir estão apresentados os dados referentes aos números de itens elaborados, aceitos e descartados para cada curso avaliado na edição do Enade de 2019.

Gráfico 1: Número de Itens elaborados, aceitos e descartados por curso Enade 2019



Fonte: Sousa (2021b, p.16).

Pode-se perceber, pelos dados apresentados no gráfico, que para alguns cursos o percentual de itens reprovados é acima de 60%, como é o caso, por exemplo, dos cursos: Curso Superior de Tecnologia em Gestão Hospitalar (62%), Enfermagem (66%) e Engenharia de Computação (64%). Este fato é uma das evidências de que há muito que se discutir, aprender, em termos de técnicas pedagógicas na construção de itens para avaliação do tipo em larga escala.

Nesse contexto, seria necessário ligar,

O alerta sinaliza para dificuldades de docentes destas áreas em elaborar itens que atendam ao padrão exigido para avaliação em larga escala. Por outro lado, ratifica-se a afirmativa de que a capacitação pedagógica se faz ainda necessária no processo de elaboração e revisão de itens, não só em função de novos componentes desse grupo de docentes, mas também pela necessidade de novas investidas de intervenções pedagógicas nesse processo. Embora o BNI-Enade esteja em uma em fase de pleno abastecimento de itens, ainda não tem apresentado indícios de colapso no sentido de carência generalizada de itens, mas é preciso estar atento para essa possibilidade no futuro próximo (SOUSA, 2021b, p. 17).

De acordo com Sousa (2021b), os principais motivos pelos quais itens são descartados estão apresentados no quadro 2.

Quadro 2: Motivos de descartes de itens para a avaliação em larga escala tipo Enade

PRINCIPAIS MOTIVOS DE DESCARTES DE ITENS	
Motivo	Justificativa
Implausibilidade.	Distratores que favoreçam a resolução mediante raciocínio por exclusão.
Nível de dificuldade.	Aquém do adequado ou muito simples/óbvio para avaliar concluintes de cursos de graduação.
Abordagem de conhecimento.	Muito específico; mais próximos das linhas especializadas de pesquisa na pós-graduação.

Uso de dados estatísticos.	Informações (centrais, conceitos, abordagem teórica) defasadas, ultrapassadas ou controversas que comprometem a validade das afirmações.
Construções textuais.	Elaborada em língua estrangeira para prova em que não cabe a adoção de outro idioma.
Reprodução de item.	Idêntico ou extremamente semelhante a outro que já foi selecionado para o BNI.
Presença de regionalismos.	Falta de abrangência mínima correspondente ao caráter nacional do Exame.
Ocorrência de falhas técnicas ou fuga ao modelo proposto pelo Inep para as provas do Enade.	Item solicita a escolha de opção incorreta. Item de asserção-razão fora do padrão. Item apresenta redação sem coesão e/ou coerência entre os elementos, estando, no todo, prejudicado. Item apresenta distratores excludentes entre si. Item apresenta distratores óbvios, que não dependem de formação em curso de graduação para serem julgados. Item depende de análise de gráfico/imagem/tabela que está ausente.
Item apresenta problemas de sensibilidade, estando, no todo, prejudicado.	Item aborda o tema de forma preconceituosa. Item apresenta problemas relacionados a questões político-partidárias.
Falta de correspondência com a encomenda.	Encaminhamento de item discursivo ao invés de item de múltipla escolha/objetivo.

Fonte: Sousa (2021b, p. 18).

Tendo essas análises como referencial, foi elaborado um questionário com 5 (cinco) questões para serem respondidas por elaboradores e/ou revisores de itens para avaliação em larga escala. Para o questionário foi apresentado o seguinte título e subtítulo, respectivamente: Elaborador/Revisor de itens produzidos para avaliação em larga escala. Ao responder às questões a seguir, foi considerado que elas estão relacionadas à avaliação em larga escala.

As tabelas a seguir apresentam os resultados apontados pelos docentes respondentes. Foram escolhidas as respostas de 15 respondentes, em função dos critérios estabelecidos para respostas válidas, dentre eles: deveria ser assinalada apenas uma alternativa para cada questão; todas as questões deveriam ser respondidas; as questões deveriam ser respondidas uma única vez; o respondente deveria estar vinculado a alguma instituição de ensino.

Tabela 2: Questionário Elaborador/Revisor de Itens produzidos para avaliação em larga escala

1) Em qual instrumento avaliativo você atuou mais vezes na elaboração/revisão de Itens?	
Alternativas	Respostas (%)
Programa Internacional de Avaliação de Alunos – Pisa.	6

Exame Nacional do Ensino Médio-Enem	60
Exame Nacional de Desempenho de Estudantes – Enade.	28
Exame Nacional de Certificação de Competências de Jovens e Adultos-ENCCEJA.	0
Outro.	6
TOTAL	100

Fonte: elaborada pelo autor (2021).

Esta questão foi apresentada com o propósito apenas de identificar que o elaborador assinalasse e se identificasse com uma avaliação em larga escala, embora identificar àqueles que se tenha interagido com o Enem e o Enade é um diferencial em função das dimensões e complexidades desses dois exames. Para Andrade e Freitag (2021), essas duas avaliações perpassam “[...] a trajetória do egresso da educação superior – Enem, no ingresso, e o Enade, na conclusão – e os itens que a compõem em um recorte com foco nas habilidades de leitura, considerando a prova de Linguagens no Enem e Formação Geral do Enade” (ANDRADE; FREITAG, 2021, p. 179).

O Enade, por exemplo, possui suas especificidades e particulares que o diferencia largamente dos outros exames citados. Os dados mostram que, aproximadamente, 90% dos respondentes concentram-se nestes dois exames principais.

Tabela 3: Questionário Elaborador/Revisor de Itens produzidos para avaliação em larga escala

2) Em quantas edições você atuou como elaborador/revisor de Itens?	
Alternativas	Respostas (%)
Uma	14
Duas	40
Três	6
Quatro	20
Cinco ou mais	20
TOTAL	100

Fonte: elaborada pelo autor (2021).

Com esta questão se esperava ter em mente o grau de experiência que os respondentes têm com avaliação em larga escala. Embora os resultados apontassem para uma pulverização, pode perceber que pelo menos 80% tiveram 2 (duas) ou mais interações com mais de duas edições e 40% com mais de três edições.

Tabela 4: Questionário Elaborador/Revisor de Itens produzidos para avaliação em larga escala

3) Como você avalia sua participação no curso de capacitação para elaborador/revisor de Itens?	
Alternativas	Respostas (%)
Participo o mínimo possível.	6
Apenas respondo às perguntas propostas.	0

Participo de todas as atividades propostas.	74
Participo somente o suficiente para ser aprovado no curso.	0
Participo integralmente do curso.	20
TOTAL	100

Fonte: elaborada pelo autor (2021).

Nesta questão a expectativa foi que o respondente realizasse uma autoavaliação de sua participação no curso de capacitação para elaboradores e revisores de item para avaliação em larga escala. Observa-se que as respostas se concentram na alternativa em que o respondente considera que participa de todas as atividades propostas durante o curso de capacitação. Isso evidencia que a maioria participa ativamente do curso o que leva a entender que ele esteja apto a participar da elaboração e revisão de itens para avaliação em larga escala.

Tabela 5: Questionário Elaborador/Revisor de Itens produzidos para avaliação em larga escala

4) Em sua opinião, qual tipo de item é mais difícil de ser elaborado/revisado?	
Alternativas	Respostas (%)
Discursivo.	20
Asserção-razão.	20
Resposta única.	40
Resposta múltipla.	14
Complementação e/ou interpretação.	6
TOTAL	100

Fonte: elaborada pelo autor (2021).

Nesta questão o propósito era identificar se havia um determinado tipo de item em que os respondentes apontassem como o mais complexo de ser elaborado ou revisado para avaliação em larga escala. Devido à dispersão das respostas dos respondentes não é seguro afirmar que há um determinado item com maior grau de dificuldade em ser elaborado ou revisado, embora o tipo de item com resposta única tenha sido apontado com 40% dos docentes respondentes.

Tabela 6: Questionário Elaborador/Revisor de Itens produzidos para avaliação em larga escala.

5) Em sua opinião, o descarte de um item está diretamente relacionado a qual fator?	
Alternativas	Respostas (%)
Realização inadequada, por parte do elaborador, no curso de capacitação.	28
Falha no curso de capacitação disponibilizado para elaboradores de Itens.	6
Falta de experiência do elaborador na confecção de Itens para esse tipo de avaliação.	40
Tipos de Itens para esse modelo de avaliação são difíceis de serem elaborados.	20

Falha na formação docente implicando em dificuldade na elaboração de Itens para esse modelo de avaliação.	6
TOTAL	100

Fonte: elaborada pelo autor (2021).

O propósito com esta questão foi avaliar as dificuldades enfrentadas pelo elaborador e/ou revisor com as metodologias utilizadas na construção desses tipos de itens. As respostas apontaram evidências de que possíveis falhas estão nas dificuldades encontradas para a elaboração desses tipos de itens e não nos processos pedagógicos do curso de capacitação disponibilizado pelo Inep. Outra evidência importante que foi detectada, está no fato de que os itens elaborados pelos docentes em seu cotidiano acadêmico, diferenciam-se dos tipos de itens de avaliação em larga escala. Isso pode ser um fator a ser observado e trabalhado na revisão da prática pedagógica desses docentes (SOUSA, 2021a).

Uma sugestão seria a implementação do curso de capacitação para elaboradores e revisores de itens no formato contínuo. Essa medida poderia contribuir com o abastecimento de itens, principalmente, cursos e/ou áreas que apresentem baixa demanda de itens armazenados em seus respectivos bancos. Além desse propósito, atenderia também à agenda de colaboradores que têm suas participações nos processos de elaboração de itens, prejudicadas em função de suas atividades desenvolvidas em suas respectivas instituições de ensino (SOUSA, 2021a).

Por fim, seria oportuno a criação de um canal que possibilitasse grupos, tanto da educação básica quanto da educação superior, para participarem como elaborador e/ou revisor de itens para avaliação em larga escala. Isso poderia ser em formatos de encontros nacionais, simpósios regionais, *workshops locais*, *webinários*, dentre outros. Nesse evento seriam disponibilizados esclarecimentos e orientações dos procedimentos necessários para a construção de itens de acordo com o padrão exigido para avaliação em larga escala. Com isso, poder-se-ia atrair mais participantes nos processos de elaboração de itens para o banco de itens da educação básica e para o banco de itens da educação superior.

5 Considerações finais

Na elaboração deste estudo, a disseminação de dados oficiais e o questionário, via *Google forms*, possibilitou uma análise mista. Os resultados apontaram para as dificuldades na elaboração e revisão de itens para avaliação em larga escala. Foi possível identificar a problemática da eficiência da produção de itens para esse tipo de avaliação.

Com isso, o que levou aos questionamentos: quais desafios são postos na produção de itens para composição das avaliações em larga escala? Que tipos de dificuldades são encontrados na elaboração de itens de qualidade para essas avaliações? Pôde-se averiguar a partir das análises de dados disponibilizados pelo Inep e obtidos por meio do questionário. Tendo como uma possível resposta a falta de experiência, por parte dos colaboradores, na elaboração de itens para avaliação em larga escala.

Dentro dessa problemática, foi possível atingir o principal objetivo desse estudo que consistiu em: verificar os desafios e as dificuldades para produção de

item de qualidade para seu efetivo aproveitamento e utilização em avaliações institucionais ou para compor seus respectivos bancos de itens.

Nesse contexto de avaliar o sistema educacional brasileiro com atributos voltados para a melhoria desses processos, pode-se inferir que às avaliações em larga escala “[...] restaram por representar uma forma de constatar, averiguar, apontar a situação da educação nacional [...] amalhando indicadores que servissem como base para a consecução do binômio acesso/sucesso” (ESQUINSANI, 2019, p. 633). Portanto, as avaliações em larga escala, como Enem e Enade, por exemplo, formam-se como uma contingência da realidade educacional vigente. Manfio (2013) citando Fischer (2010) defende que “[...] a finalidade crucial da avaliação em larga escala é tomar decisões pedagógicas e administrativas [...] estão vinculadas as políticas educacionais a partir do projeto de globalização” (MANFIO, p. 31334).

Portanto, percebe-se que a avaliação em larga escala pode ser considerada uma fotografia da realidade escolar, um indicador, ou mesmo uma representação mais aproximada possível das competências e habilidades que os estudantes possuem durante sua jornada estudantil, uma comparação que possui fundamento científico que certamente traz indicativos importantes para o planejamento pedagógico (SOUZA, 2019).

Por outro lado, estudos contemporâneos sobre aprendizagem e pensamento iluminam construções críticas que permanecem inexploradas pela maioria das avaliações em larga escala, incluindo a forma como os estudantes organizam os conhecimentos, as representações que geram para resolver problemas, a sua utilização de estratégias e capacidades de autocontrole, e as suas contribuições individuais para a resolução de problemas de grupo (CHUDOWSKY; PELLEGRINO, 2003).

Este estudo esteve nos limites de dados apresentados pelo Inep, órgão responsável pelos bancos de itens. Nesse sentido, podem surgir dificuldades na comparação com outras fontes de dados, como por exemplo, as variáveis analisadas neste estudo, que foram definidas de modo arbitrário conforme a conveniência deste estudo, como taxa de Aproveitamento de itens do BNI-Enade, por exemplo.

Com isso, outras pesquisas poderiam confrontar os dados deste estudo ou mesmo realizar a comparabilidade da produção de itens em outras edições dos exames do Inep; o custo efetivo financeiro da gestão do BNI-Enade e, assim, analisar a demanda de Itens aceitos (aprovados) para esses bancos de itens. Além disso, apontar outros meios que venham a contribuir com a qualidade na produção de itens de para compor os bancos de itens do Inep.

6 Referências

ALAVARSE, Ocimar Munhoz. **Avaliar as avaliações em larga escala: desafios políticos**. 12 dez. 2013. Disponível em: <https://revistaeducacao.com.br/2013/12/12/avaliar-as-avaliacoes-em-larga-escala-desafios-politicos/>. Acesso em: 10 maio 2021.

AMARO, Ivan. Avaliação em larga escala e qualidade: dos enquadres regulatórios aos caminhos alternativos. **Linhas críticas**, v.22, n.48. Brasília/DF, maio/ago. 2016.

BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela. Avaliações em larga escala: uma sistematização do debate. **Educação pesquisa**, v. 41, n. especial. São Paulo, 2015.

BRASIL. Controladoria Geral da União. **Plataforma integrada de ouvidoria e acesso à informação**. Disponível em: <https://falabr.cgu.gov.br/publico/Manifestacao/SelecionarTipoManifestacao.aspx?ReturnUrl=%2f>. Acesso em: 20 de maio 2021.

BRASIL. Planalto. **Constituição da República Federativa do Brasil de 1988**. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 25 de jun. 2021.

BRASIL. Planalto. **Lei nº 9.394, de 20 de dezembro de 1996**. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 23 de jun. 2021.

CAMPOS, Fernanda Cristina dos Santos. **Elaboração da prova do ENADE e no modelo do banco nacional de itens**. Dissertação (Mestrado – Programa de pós-graduação profissional em gestão e avaliação da educação pública. Área de concentração: Centro de Políticas Públicas e Avaliação da Educação) Faculdade de Educação - CAEd/UFJF, Juiz de Fora/MG, 2013.

CERRI, Rosilene. Palestra na PUC-Campinas: **BNI-ENADE**. Disponível em: <https://www.puc-campinas.edu.br/wp-content/uploads/2016/04/proavi-palestra-sobre-bni-seminarios-enade2014.pdf>. Acesso em: 13 de maio 2021.

CHUDOWSKY Naomi; PELLEGRINO, James. Large-Scale Assessments That Support Learning: What Will It Take? **Theory into practice**, v. 42, n. 1, College of Education, The Ohio State University, 2003.

ESQUINSANI, Rosimar Serena Siqueira. Os anos 1980 e as avaliações em larga escala como fenômeno histórico. **El futuro del pasado**: revista electrónica de historia, n. 10, Espanha, 2019.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais – Anísio Teixeira. Guia BNI: Atualizado em 19 de fevereiro de 2020. Disponível em: <http://inep.gov.br/guia-bni1>. Acesso em: 10 de maio 2021.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais – Anísio Teixeira. **O que é o BNI**. Disponível em: <http://inep.gov.br/banco-nacional-de-itens>. Acesso em: 11 de jun. 2021.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais – Anísio Teixeira. **O que é o ENADE**. Disponível em: <http://portal.inep.gov.br/enade>. Acesso em: 11 de maio 2021b.

INEP. Instituto Nacional de Estudos e Pesquisas Educacionais – Anísio Teixeira. **Produção de itens para prova ENADE**. Manifestação Protocolo: nº 23546.025418/2021-09. Brasília/DF, 2021c.

LIMA, Marcos Antônio Martins; MAIA, José Leudo; CIASCA, Maria Isabel Filgueiras Lima; SOUZA, Jacqueline Ramos Macedo Antunes. Avaliação da Educação Superior no Brasil: análise do Índice Geral dos Cursos (IGC) numa perspectiva quali/quantitativa. **Avaliação**, v. 25, n. 03, Campinas; Sorocaba, 2020.

MANFIO, Aline. Avaliação em larga escala e qualidade de ensino: análise da produção em periódicos qualificados (1995-2012). **XI Congresso Nacional de Educação**, PUC/PR, Curitiba, set. 2013.

SOUZA, Atair Silva. Um breve panorama da produção de itens para avaliação da qualidade da Educação Superior em 2019. **Revista relações sociais (REVES)**, v.4, n.4. Viçosa/MG, 2021. p. 1-12

SOUSA, Atair Silva. Um panorama da qualidade dos itens elaborados para o banco nacional de itens da educação superior na edição do Enade de 2019. **Revista exitus**, v. 11, e020167. Santarém/PA, 2021. p. 01-23

SOUSA, Clarilza Prado; FERREIRA, Sandra Lúcia. Avaliação de larga escala e da aprendizagem na escola: um diálogo necessário. **Psicologia da educação**, 48. PUC/SP. São Paulo, 2019.

SOUSA, Sandra Zákia; ARCAS, Paulo Henrique. Implicações da Avaliação em Larga Escala no Currículo: revelações de escolas estaduais de São Paulo. **Educação: Teoria e Prática**, v. 20, n.35. São Paulo, jul./dez. 2010.

SOUZA, José Eduardo Pereira. Avaliação em larga escala, o senso comum, críticas e ponderações. **Revista Do Instituto de Políticas Públicas de Marília**, v.5, n.2, Marília/SP, jul./dez. 2019.

WELTER, Cristiane Backes; PASINI, Juliana Fatima Serraglio. Avaliação em larga escala: que avaliação? O que dizem as crianças? **X Anped Sul**, Florianópolis, out. 2014.

WELTER, Cristiane Backes; WERLE, Flávia Obino Corrêa. Processos de invisibilização na avaliação em larga escala. **Ensaio: avaliação e políticas públicas em educação**, v.29, n.111, Rio de Janeiro, abr./jun. 2021.

WIEBUSCH, Eloisa Maria. Avaliação em larga escala: uma possibilidade para a melhoria da aprendizagem. **IX Anped Sul. Seminário de pesquisa em educação da região sul**. PUC/RS, Porto Alegre, 2012.

Informações complementares

Financiamento

Não se aplica.

Contribuição de autoria

Concepção e elaboração do manuscrito: Atair Silva de Sousa.

Coleta de dados: Atair Silva de Sousa.

Análise de dados: Atair Silva de Sousa.

Discussão dos resultados: N Atair Silva de Sousa.

Revisão e aprovação: Atair Silva de Sousa.

Preprint, originalidade e ineditismo

O artigo é original, inédito e não foi depositado como *preprint*.

Verificação de similaridades

O artigo foi submetido ao iThenticate, em 15 de novembro de 2023, e obteve um índice de similaridade compatível com a política antiplágio da revista Pesquisa e Debate em Educação.

Consentimento de uso de imagem

Não se aplica.

Aprovação de Comitê de Ética em Pesquisa

Não se aplica.

Conflito de interesse

Não há conflitos de interesse.

Conjunto de dados de pesquisa

Não há dados disponibilizados. Os dados podem ser disponibilizados em bancos ou repositórios próprios para tal finalidade ou, até mesmo, na nuvem (com acesso público).

Licença de uso

Os autores cedem à Revista Pesquisa e Debate em Educação os direitos exclusivos de primeira publicação, com o trabalho simultaneamente licenciado sob a [Licença Creative Commons Attribution \(CC BY\) 4.0 International](#). Esta licença permite que terceiros remixem, adaptem e criem a partir do trabalho publicado, atribuindo o devido crédito de autoria e publicação inicial neste periódico. Os autores têm autorização para assumir contratos adicionais separadamente, para distribuição não exclusiva da versão do trabalho publicada neste periódico (ex.: publicar em repositório institucional, em site pessoal, publicar uma tradução, ou como capítulo de livro), com reconhecimento de autoria e publicação inicial neste periódico.

Publisher

Universidade Federal de Juiz de Fora (UFJF), Faculdade de Educação (FACED), Centro de Políticas Públicas e Avaliação da Educação (CAEd), Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública (PPGP). Publicação no Portal de Periódicos da UFJF. As ideias expressadas neste artigo são de responsabilidade de seus autores, não representando, necessariamente, a opinião dos editores ou da universidade.

Editores

Frederico Braida; Liamara Scortegagna; Wagner Silveira Rezende.

Formato de avaliação por pares

Revisão duplamente cega (*Double blind peer review*).

Sobre os autores

Atair Silva de Sousa

Graduado - Licenciado em Física (UFMG). Especialista em Ensino de Ciências Física (UFMG). Especialista em Gestão Pública (UFMG). Mestrado em Gestão e Avaliação da Educação Pública (UFJF). Doutorando em Educação (USP)
Currículo Lattes: <http://lattes.cnpq.br/2459273724635926>