Revista Ibero-Americana de Humanidades, Ciências e Educação — REASE

doi.org/10.51891/rease.v12i11.32846

BANCO DE QUESTÕES E ELABORAÇÃO DE AVALIAÇÕES NO ENSINO SUPERIOR: UM ESTUDO NO CEST

QUESTION BANKS AND EXAM PREPARATION IN HIGHER EDUCATION: A STUDY AT CEST

Gustavo Santos Leite[1]
Thiago Nelson Faria dos Reis[2]

RESUMO: Este estudo analisou em que medida o banco de questões integrado ao PROINT influenciou a elaboração de avaliações por docentes do Centro Universitário Santa Terezinha (CEST). A pesquisa aplicada, de abordagem mista, fundamentou-se na Design Science Research e avaliou um módulo desenvolvido anteriormente. O questionário foi aplicado antes e após o uso da ferramenta na elaboração de provas reais, reunindo 12 respondentes no primeiro momento e dez no segundo, com dez participantes pareados. Os dados foram examinados por contagens, percentuais, comparação de categorias de tempo e leitura temática das respostas abertas. Todos os respondentes da etapa posterior declararam cadastrar questões novas, e oito apontaram o tempo de cadastro inicial como dificuldade. Cinco perceberam pequena redução do tempo, com divergências entre essa percepção e as faixas informadas. Quatro docentes relataram melhora no processo, três não perceberam diferença significativa e três relataram maior dificuldade. Os resultados indicaram contribuições operacionais percebidas e necessidades de aperfeiçoamento da interface e do cadastro. O desenho não permitiu quantificar economia de tempo atribuível ao módulo ou demonstrar benefícios acumulados de reutilização em ciclos posteriores.

Palavras-chave: Bancos de questões. Avaliação da aprendizagem. Ensino superior.

ABSTRACT: This study examined how the question bank integrated into PROINT influenced exam preparation among faculty members at Centro Universitário Santa Terezinha (CEST). This applied, mixed-methods study was grounded in Design Science Research and evaluated a previously developed module. A questionnaire was administered before and after use of the tool to prepare actual exams, with 12 respondents initially and ten subsequently, yielding ten matched participants. Analysis used counts, percentages, comparison of time categories, and thematic reading of open-ended responses. All respondents at the second time point reported entering new questions, and eight identified the time required for initial question entry as a difficulty. Five perceived a slight reduction in preparation time, although this perception did not always align with their reported time categories. Four faculty members reported improvements in the process, three perceived no significant difference, and three reported greater difficulty. The findings indicated perceived operational benefits and a need to improve the interface and question-entry process. The design did not allow time savings attributable to the module to be quantified or cumulative benefits from question reuse in subsequent cycles to be demonstrated.

Keywords: Question banks. Learning assessment. Higher education.

[1] Graduando em Sistemas de Informação. Centro Universitário Santa Terezinha (CEST).
[2] Doutor. Pós-Doutor pela Universidade Federal do Maranhão (UFMA). Professor. Orientador. Centro Universitário Santa Terezinha (CEST).

1 INTRODUÇÃO

A elaboração de avaliações no ensino superior envolve decisões sobre objetivos de aprendizagem, conteúdos e questões, tomadas conforme as condições em que os professores trabalham. Ao discutirem o desenho da avaliação em um mundo digital, Bearman, Nieminen e Ajjawi (2023) reforçam que as escolhas avaliativas e tecnológicas são influenciadas pelo contexto de utilização. Investigar uma ferramenta de apoio à elaboração de provas exige, assim, compreender como seus recursos se articulam com as atividades dos docentes.

As atividades de preparação do ensino também integram o trabalho docente. Kenny e Fluck (2023), em entrevistas com acadêmicos de universidades australianas, identificaram que tarefas de planejamento e preparação nem sempre eram adequadamente reconhecidas pelos modelos institucionais de distribuição da carga de trabalho. Embora realizado em contexto institucional distinto, o estudo contribui para discutir a importância dessas atividades no conjunto das atribuições dos professores. No caso das provas, a preparação inclui produzir ou selecionar questões e organizar o documento que será aplicado.

Os bancos de questões estão entre os recursos utilizados nessa preparação, pois permitem armazenar, classificar e selecionar itens para diferentes avaliações. Krzic e Brown (2022) descrevem sua utilização em uma disciplina introdutória de Ciência do Solo, com ênfase na organização das questões por tema e nível de dificuldade e no alinhamento aos objetivos de aprendizagem. Em uma disciplina de Química, Beerepoot (2023) apresenta aplicações em avaliações formativas e somativas e destaca a revisão e o acompanhamento da qualidade dos itens. A utilização desses recursos também demanda elaboração, organização e revisão do conteúdo avaliativo.

Parte das vantagens relatadas nesses estudos está associada à aplicação de testes digitais e à correção automatizada. Embora descrevam aplicações e possibilidades dos bancos de questões, os trabalhos de Krzic e Brown (2022) e Beerepoot (2023) não comparam o processo de elaboração de provas pelos mesmos docentes antes e depois da adoção de uma ferramenta institucional. Para compreender essa relação em um contexto específico, é necessário examinar as atividades apoiadas pelo sistema e a experiência dos professores que o utilizam.

No Centro Universitário Santa Terezinha (CEST), antes da disponibilização do módulo investigado, a elaboração das provas acadêmicas seguia um fluxo de edição manual. A coordenação fornecia um modelo, que os professores preenchiam em editores de texto, encaminhavam para aprovação e imprimiam por conta própria. O PROINT, plataforma

desenvolvida pela Coordenação de Inovação e Tecnologia (CIT) e em operação desde 2024, passou a incorporar recursos para cadastro e organização de questões, montagem de avaliações e geração de provas em formato de documento portátil (PDF). Essas funcionalidades foram desenvolvidas com a finalidade de centralizar etapas e facilitar a preparação dos instrumentos avaliativos. Sua disponibilidade, entretanto, não permite concluir, por si só, quais mudanças ocorreram no trabalho docente. O presente estudo concentra-se no cadastro e na seleção de questões, na montagem das avaliações e na geração do documento formatado em PDF, sem abranger a correção das provas.

Diante desse contexto, este artigo tem como objetivo analisar em que medida a utilização de um banco de questões integrado ao sistema PROINT influencia o processo de elaboração de avaliações pelos docentes do CEST. Trata-se de uma pesquisa aplicada, de abordagem mista, com questionários referentes ao processo anterior e à experiência posterior ao uso do módulo. A investigação busca produzir evidências contextualizadas sobre a utilização da ferramenta e subsidiar seu aprimoramento institucional.

O artigo apresenta o referencial sobre prática avaliativa, bancos de questões e avaliação de artefatos, seguido dos procedimentos metodológicos, dos resultados e de sua discussão. As considerações finais retomam as contribuições e os limites da investigação.

2 REFERENCIAL TEÓRICO

2.1 Elaboração de avaliações como prática docente

A prática avaliativa compreende as decisões e os procedimentos pelos quais os professores acompanham e interpretam a aprendizagem dos estudantes. A escolha de uma prova, de um projeto ou de uma atividade em grupo integra esse processo, mas não esgota suas possibilidades. Em revisão de 52 artigos, Pezzi, Veiga e Fernandes (2025) identificaram a predominância de provas, testes e exames no conjunto de estudos analisado, ao lado de práticas como portfólios, autoavaliação e avaliação por pares. Os autores defendem que a seleção dos instrumentos seja orientada pelos objetivos de aprendizagem e pela diversidade de conhecimentos e habilidades que se pretende avaliar.

As escolhas dos docentes também se relacionam com suas trajetórias formativas. Gonzaga et al. (2023), em pesquisa qualitativa com 22 professores de cursos de Licenciatura em Matemática dos institutos federais do Ceará e da Paraíba, identificaram limitações na formação específica sobre avaliação da aprendizagem. Os relatos indicaram a importância das

experiências anteriores, tanto como estudantes quanto como professores, na constituição dos saberes avaliativos. No contexto investigado, parte das práticas se apoiava na reprodução de modelos conhecidos e em decisões intuitivas. O estudo permite compreender como formação e experiência se articulam nas práticas desses participantes, sem autorizar a generalização desse perfil para todo o ensino superior.

A preparação dos instrumentos avaliativos integra esse conjunto de decisões. Selecionar conteúdos e questões exige considerar o que se pretende avaliar e como a atividade proposta permitirá obter informações sobre a aprendizagem. Bearman, Nieminen e Ajjawi (2023) adotam uma concepção de desenho da avaliação que abrange as tarefas, os processos de atribuição de notas, o feedback e as relações entre esses elementos. Sob essa perspectiva, preparar uma prova envolve mais do que reunir questões em um documento: sua composição precisa ser considerada em relação às finalidades da avaliação e às condições em que será utilizada.

Quando recursos digitais passam a apoiar esse trabalho, os conhecimentos pedagógicos permanecem relevantes. Viberg et al. (2024) investigaram conhecimentos e práticas de avaliação digital declarados por 219 docentes de uma universidade europeia. No modelo analisado, os conhecimentos pedagógicos e pedagógicos do conteúdo apresentaram relações com diferentes dimensões das práticas de avaliação digital. Os resultados reforçam a importância de apoiar os professores no desenvolvimento dessas práticas, embora o caráter transversal e o uso de autorrelato não permitam estabelecer relações causais. A disponibilidade de uma ferramenta, portanto, precisa ser examinada em conjunto com as decisões e os conhecimentos de quem a utiliza.

Para analisar o apoio tecnológico à elaboração de provas, convém distinguir as decisões pedagógicas das operações necessárias à preparação do instrumento. As primeiras dizem respeito, por exemplo, à seleção do conteúdo e à adequação das questões aos objetivos pretendidos. As operações incluem armazenar e recuperar itens, ordenar as questões e formatar o documento. Essa distinção é adotada neste estudo como recurso analítico, reconhecendo que ambas se articulam durante o trabalho docente. Uma alteração na maneira de organizar ou formatar a prova pode modificar sua preparação, sem demonstrar, por si só, melhoria na qualidade pedagógica da avaliação.

O banco de questões pode ser examinado a partir dessa articulação. Seus recursos oferecem suporte ao armazenamento, à recuperação e à seleção de itens, enquanto a definição do que será avaliado e a apreciação da pertinência das questões continuam exigindo decisões

docentes. Neste estudo, interessa analisar como esse suporte participa da elaboração das avaliações, considerando as atividades realizadas e as percepções dos professores. A discussão sobre a organização e a utilização dos bancos de questões aprofunda essa relação na subseção seguinte.

2.2 Bancos de questões no ensino superior

Nos estudos considerados aqui, o banco de questões aparece como um conjunto organizado de itens, classificados e recuperáveis para compor diferentes avaliações. Os critérios de organização variam. Krzic e Brown (2022) descrevem bancos usados em uma disciplina introdutória de Ciência do Solo, organizados por tema e nível de dificuldade e alinhados aos objetivos de aprendizagem. Beerepoot (2023) relata mais de 180 bancos associados aos resultados de aprendizagem de uma disciplina de Química, empregados em avaliações formativas e somativas.

A qualidade dos itens recebe atenção nessas experiências. Beerepoot (2023) destaca a revisão e a garantia de qualidade e descreve como trabalhosa a construção de bancos compostos por itens de qualidade. Krzic e Brown (2022) apontam, entre os desafios, a necessidade de produzir múltiplas versões das questões, preparar orientações de correção e selecionar os itens com cuidado. Os relatos indicam que o uso de um banco envolve também elaboração, revisão e manutenção do conteúdo.

A construção pode ser coletiva. Namesnik-Silvester et al. (2025) apresentam o ShareStats, banco aberto de itens de Estatística desenvolvido por uma comunidade de docentes do ensino superior. O projeto articula uma taxonomia de classificação a um procedimento de revisão que envolve mais de um integrante da equipe, além da padronização dos arquivos para reutilização em diferentes ambientes digitais. A experiência mostra como o compartilhamento de questões pode ser acompanhado por procedimentos de organização e revisão do conteúdo. O estudo, contudo, não compara o tempo de elaboração das provas antes e depois da utilização do banco.

No contexto brasileiro, Barbosa, Silva e Freitas (2022) descrevem um sistema web de banco de questões desenvolvido para o Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), no qual os professores cadastram e classificam questões e a inclusão no banco depende da validação por outros docentes. O sistema foi avaliado por nove professores do campus Cedro, por meio de questionário aplicado após um período de uma semana

disponibilizado para utilização e testes. A maioria dos participantes concordou com as afirmações apresentadas sobre as funcionalidades. Esse período, contudo, não indica uso contínuo do sistema, e o estudo não demonstra redução mensurada do tempo de elaboração.

As experiências discutidas relacionam os bancos de questões à organização, à revisão e à reutilização de itens, mas diferem quanto às finalidades e às formas de avaliação das ferramentas. Descrever funcionalidades e registrar sua aceitação pelos usuários oferece informações relevantes, porém não esclarece, por si só, como o processo de elaboração das avaliações se modifica. Neste estudo, essa relação é examinada pela comparação dos relatos docentes referentes aos momentos anterior e posterior ao uso do módulo, considerando os limites do instrumento e do contexto investigado.

2.3 Avaliação de artefatos tecnológicos no contexto educacional

A investigação de uma tecnologia institucional exige relacionar o problema que motivou seu desenvolvimento, os objetivos da solução e as evidências obtidas em sua utilização. A Design Science Research (DSR) oferece uma abordagem para essa articulação, ao orientar pesquisas que desenvolvem e avaliam artefatos destinados à resolução de problemas. Peffers et al. (2007) propõem uma metodologia composta por seis atividades: identificação do problema e motivação; definição dos objetivos da solução; projeto e desenvolvimento; demonstração; avaliação; e comunicação. O modelo admite diferentes pontos de entrada, conforme as condições que deram origem à investigação.

Nesse modelo, demonstração e avaliação cumprem funções distintas. A demonstração apresenta o uso do artefato para resolver situações relacionadas ao problema, enquanto a avaliação examina em que medida os resultados observados correspondem aos objetivos da solução. Peffers et al. (2007) apresentam diferentes possibilidades de avaliação, incluindo medidas de desempenho, pesquisas de satisfação e manifestações dos usuários. A escolha dos procedimentos depende das características do artefato, do contexto de utilização e dos objetivos investigados.

No campo educacional brasileiro, Galvão, Madureira e Schneider (2024) analisaram 76 teses e dissertações, publicadas entre 2019 e 2023, que empregaram a DSR na produção de artefatos voltados à educação. A revisão identificou diferentes tipos de produtos, protocolos e procedimentos avaliativos, com destaque para a utilização de questionários na coleta de dados. Os autores também observaram adaptações das etapas metodológicas e ressaltaram a

necessidade de justificar essas escolhas e explicitar os procedimentos realizados. O estudo evidencia a diversidade de aplicações da DSR na educação e a importância da coerência entre o problema investigado e as estratégias adotadas para avaliar a solução proposta.

Neste estudo, o módulo de banco de questões do PROINT é examinado como um artefato de apoio à elaboração de provas. A comparação dos relatos docentes referentes aos momentos anterior e posterior ao uso relaciona os objetivos desse suporte às mudanças percebidas, às dificuldades e às condições de utilização. Os procedimentos dessa avaliação são apresentados na seção seguinte.

3 METODOLOGIA

3.1 Delineamento e contexto

A pesquisa, de natureza aplicada, adotou abordagem mista e caráter descritivo e analítico. Foi realizada no Centro Universitário Santa Terezinha (CEST), instituição privada de ensino superior localizada em São Luís, Maranhão, e examinou o processo de elaboração de avaliações por docentes antes e após a utilização do módulo de banco de questões integrado ao PROINT. A investigação ocorreu no ambiente de trabalho dos participantes, sem distribuição aleatória ou grupo de controle.

A avaliação do artefato foi fundamentada na DSR, relacionando os objetivos da solução às evidências obtidas junto aos usuários, conforme Peffers et al. (2007). O desenvolvimento antecedeu a pesquisa; por isso, o estudo concentrou-se na avaliação da utilização, sem pressupor adoção integral do protocolo nas atividades anteriores.

3.2 Artefato e recorte da investigação

O módulo permite cadastrar e organizar questões, recuperar itens por filtros, montar avaliações e gerar o documento da prova em PDF. A investigação focalizou o apoio dessas funcionalidades à preparação das provas pelos docentes.

O módulo foi disponibilizado em ambiente de produção em 3 de agosto de 2026. O recorte compreendeu a elaboração de avaliações reais referentes à primeira prova do segundo semestre de 2026, denominada institucionalmente NP1, cuja aplicação ocorreu em 23 de setembro. As respostas posteriores ao uso foram fornecidas após a elaboração das provas, podendo anteceder sua aplicação aos estudantes. O período correspondeu ao uso inicial do módulo e à alimentação do banco de questões.

3.3 Participantes e coleta de dados

Foram convidados 13 docentes do CEST para os dois momentos da pesquisa. A participação foi voluntária, sem seleção probabilística. Os critérios de participação contemplaram docentes que ministrassem disciplinas com avaliações objetivas gerenciadas pelo PROINT e utilizassem o módulo na elaboração de avaliações reais no período investigado.

A coleta foi realizada pelo Google Forms, de forma digital e assíncrona. O bloco A foi disponibilizado em 1º de julho de 2026, para preenchimento antes da primeira utilização do módulo por cada participante. Os blocos B e C foram disponibilizados conjuntamente em 25 de agosto, para preenchimento após a elaboração de uma avaliação com a ferramenta. O instrumento não foi alterado durante a aplicação.

A base considerada nesta análise reuniu 12 respostas docentes no bloco A e dez nos blocos B e C, com dez participantes pareados. As respostas posteriores incluídas foram registradas entre 31 de agosto e 10 de novembro de 2026. Os dois docentes sem resposta no segundo momento integraram somente a descrição do processo anterior. O quantitativo incluído atingiu a meta de ao menos dez respostas por bloco prevista no projeto.

3.4 Instrumento

O questionário aplicado compreendeu 16 perguntas principais distribuídas em três blocos, além de campos complementares. O bloco A reuniu cinco perguntas sobre tempo habitual de elaboração, etapas realizadas, reaproveitamento de questões, dificuldades e compatibilidade do tempo despendido com a carga de trabalho. O bloco B apresentou quatro perguntas sobre tempo de elaboração com o módulo, tamanho da avaliação, reutilização de questões e frequência de uso das funcionalidades. O bloco C reuniu sete perguntas sobre mudanças percebidas no processo, impacto percebido no tempo, aspectos úteis, dificuldades, intenção de continuidade, avaliação geral e sugestões ou comentários.

Foram utilizadas alternativas fechadas, seleção de múltiplas opções, escalas de cinco pontos e campos abertos. Os tempos, o tamanho das avaliações e a proporção de questões reutilizadas foram informados em categorias. As informações decorreram de autorrelato, sem cronometragem ou verificação independente por registros do sistema.

Na transposição do instrumento para o Google Forms, foram adotadas configurações distintas das apresentadas no questionário anexado ao protocolo. A questão 4 permitiu múltiplas marcações, contabilizadas separadamente na análise, sem atribuir a cada participante

uma única dificuldade principal. Na questão 9, as frequências foram dispostas nas linhas e as funcionalidades nas colunas, invertendo a disposição do questionário anexado ao protocolo. Essa configuração não assegurou uma resposta para cada funcionalidade. As omissões foram preservadas, sem interpretá-las como ausência de utilização.

O instrumento foi elaborado a partir dos objetivos da investigação e das funcionalidades previstas para o módulo de banco de questões do PROINT. As perguntas buscaram registrar aspectos do processo de elaboração das avaliações, incluindo tempo despendido, etapas realizadas, dificuldades percebidas, reutilização de questões, frequência de uso das funcionalidades e intenção de continuidade. As questões abertas foram incluídas para possibilitar que os participantes registrassem percepções, sugestões e dificuldades não contempladas integralmente pelas alternativas fechadas.

Antes da aplicação, o instrumento foi submetido à apreciação do orientador e da equipe da Coordenação de Inovação e Tecnologia (CIT), com o objetivo de verificar a clareza da redação, a pertinência das perguntas e sua relação com os objetivos da pesquisa.

3.5 Preparação da base e análise

Os arquivos originais foram preservados, e a base de análise foi organizada com códigos individuais mantidos entre os momentos. O pareamento foi estabelecido por coincidência de e-mail em oito casos, por correspondência entre endereços confirmada pelo autor em um caso e pela identificação, também confirmada pelo autor, de uma resposta sem e-mail em outro. Esta última foi recuperada por conferência visual do PDF de respostas do bloco A.

Dos 15 registros disponíveis no PDF do bloco A, um foi excluído por corresponder a um teste do pesquisador. Outro, cuja identidade e elegibilidade não puderam ser confirmadas, foi separado da análise, permanecendo preservado no documento original. Um terceiro registro não foi incluído por não apresentar TCLE assinado até a atualização da base. As decisões de inclusão, exclusão, recuperação e pareamento foram registradas para permitir sua revisão.

A análise quantitativa empregou contagens e percentuais, com denominadores correspondentes às respostas válidas de cada pergunta. Nas questões de múltipla seleção, cada alternativa foi contabilizada separadamente. Respostas ausentes não foram imputadas, e combinações ambíguas foram preservadas. A descrição do processo anterior considerou os 12 participantes, enquanto a comparação individual entre os momentos utilizou somente os dez pares.

As faixas de tempo diferiram entre os formulários e foram apresentadas em suas categorias originais. Não foram atribuídos pontos médios para estimar minutos ou calcular percentuais de redução de tempo. A comparação das categorias foi examinada em conjunto com a pergunta específica sobre impacto percebido, considerando que o bloco A descreveu a rotina habitual e o segundo momento se referiu à avaliação elaborada com o módulo.

As respostas abertas foram submetidas a leitura integral e repetida. Inicialmente, foram identificados trechos relacionados a dificuldades, contribuições percebidas, sugestões de aprimoramento e condições de uso do módulo. Em seguida, esses trechos foram agrupados em eixos temáticos provisórios, posteriormente revistos em comparação com as perguntas fechadas e com os objetivos da pesquisa. Os eixos finais foram: interface e fluxo de trabalho; esforço de cadastro e retrabalho; organização e institucionalização; e condições de utilidade.

Os dados foram organizados em planilha, os cálculos foram conferidos com scripts reproduzíveis e os gráficos foram produzidos com Matplotlib.

3.6 Aspectos éticos e posição do pesquisador

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa do Centro Universitário Santa Terezinha (CEST), sob o parecer nº 8.545.348, de 27 de junho de 2026, CAAE nº 97734926.7.0000.0454. O consentimento foi registrado antes das respostas, mediante assinatura digital do Termo de Consentimento Livre e Esclarecido, disponibilizado em PDF. A participação foi voluntária, com possibilidade de desistência sem prejuízo institucional.

O autor participou do desenvolvimento do módulo investigado. Essa posição foi considerada na interpretação dos dados, por possibilitar conhecimento do contexto e também influência sobre a interação com os participantes e a leitura dos resultados. A apresentação preservou as manifestações críticas e explicitou essa condição. Os participantes foram identificados por códigos na base analítica e no artigo, mantendo-se a chave de correspondência com os e-mails separadamente. Nomes e endereços de e-mail não foram apresentados no manuscrito.

4 RESULTADOS

4.1 Participação e características do processo anterior

Dos 13 docentes convidados, foram consideradas 12 respostas no bloco A e dez nos blocos B e C. Os dez respondentes do segundo momento também participaram do primeiro,

constituindo o grupo utilizado na comparação individual. Os outros dois docentes incluídos integraram apenas a descrição do processo anterior.

No bloco A, quatro participantes informaram elaborar uma avaliação entre 30 e 60 minutos e quatro entre uma e duas horas, correspondendo a 33,3% em cada faixa. Três docentes indicaram entre duas e três horas (25,0%) e um informou mais de três horas (8,3%). Nenhum selecionou menos de 30 minutos.

A criação de questões do zero foi assinalada por 11 docentes (91,7%). A busca de questões em materiais externos, a revisão do conteúdo e a formatação do documento receberam oito marcações cada (66,7%), enquanto a adaptação de questões de semestres anteriores foi indicada por cinco participantes (41,7%). Como a pergunta admitia múltiplas respostas, as etapas podem coexistir no processo de um mesmo docente.

O reaproveitamento de questões anteriores foi descrito como raro por oito participantes (66,7%), frequente por dois (16,7%), inexistente por um (8,3%) e realizado às vezes por um (8,3%). A dificuldade mais assinalada foi manter variedade e qualidade dos itens avaliativos, com seis marcações (50,0%). A falta de tempo para criar questões e a formatação do documento receberam quatro marcações cada (33,3%). Quatro docentes também selecionaram não enfrentar dificuldades significativas (33,3%); em duas dessas respostas, a alternativa apareceu em conjunto com dificuldades específicas.

Na escala de compatibilidade do tempo de elaboração com a carga de trabalho, de 1, muito incompatível, a 5, totalmente compatível, quatro docentes atribuíram nota 3 (33,3%), três nota 2 (25,0%), três nota 5 (25,0%) e dois nota 4 (16,7%). Não houve marcação da nota 1.

4.2 Elaboração com o módulo e comparação dos participantes pareados

Os dez respondentes do segundo momento declararam que a avaliação elaborada continha entre seis e dez questões. Todos informaram cadastrar as questões como novas, sem reutilizar itens previamente disponíveis no banco naquela avaliação.

O tempo de elaboração com o módulo foi situado entre 30 e 60 minutos por quatro participantes (40,0%), entre uma e duas horas por quatro (40,0%), entre 15 e 30 minutos por um (10,0%) e abaixo de 15 minutos por um (10,0%). Nenhum selecionou mais de duas horas. As respostas individuais dos mesmos dez docentes nos dois momentos são apresentadas com o impacto percebido sobre o tempo (Tabela 1).

Tabela 1 – Faixas de tempo e impacto percebido entre os dez participantes pareados

Código	Tempo habitual anterior	Tempo com o módulo	Impacto percebido no tempo
P01	Entre 30 e 60 minutos	Menos de 15 minutos	Reduziu um pouco
P02	Entre 1 e 2 horas	Entre 30 e 60 minutos	Reduziu um pouco
P03	Entre 30 e 60 minutos	Entre 30 e 60 minutos	Reduziu um pouco
P04	Entre 2 e 3 horas	Entre 30 e 60 minutos	Aumentou um pouco
P06	Entre 2 e 3 horas	Entre 30 e 60 minutos	Não alterou
P09	Entre 1 e 2 horas	Entre 1 e 2 horas	Aumentou um pouco
P10	Entre 30 e 60 minutos	Entre 1 e 2 horas	Não alterou
P11	Mais de 3 horas	Entre 1 e 2 horas	Reduziu um pouco
P12	Entre 1 e 2 horas	Entre 1 e 2 horas	Aumentou um pouco
P13	Entre 30 e 60 minutos	Entre 15 e 30 minutos	Reduziu um pouco

Fonte: Elaborado pelos autores (2026).

Nota: A descrição habitual anterior e a experiência com a avaliação elaborada foram obtidas por autorrelato. As faixas diferem entre os instrumentos e não foram convertidas em valores médios de minutos. Os rótulos do impacto percebido foram abreviados para apresentação.

As faixas informadas e a percepção direta do impacto não foram plenamente concordantes. P04 indicou uma faixa posterior de menor duração e relatou aumento do tempo; P06 também selecionou uma faixa inferior, mas não percebeu alteração; e P10 informou uma faixa posterior de maior duração e declarou que o tempo não se alterou. P03 permaneceu na mesma faixa e percebeu pequena redução, enquanto P09 e P12 também permaneceram na mesma faixa e relataram pequeno aumento do tempo. Essas respostas foram mantidas na comparação, sem substituição de uma medida pela outra.

4.3 Mudanças percebidas e experiência de uso

As percepções sobre o processo de elaboração, o impacto percebido no tempo e a intenção de continuidade são apresentadas conjuntamente (Tabela 2).

Tabela 2 – Percepções sobre o módulo e intenção de continuidade (n = 10)

Dimensão	Resposta	n	%
Processo de elaboração	Melhorou bastante	2	20,0

Dimensão	Resposta	n	%
Processo de elaboração	Melhorou um pouco	2	20,0
Processo de elaboração	Não percebeu diferença significativa	3	30,0
Processo de elaboração	Ficou um pouco mais difícil	3	30,0
Impacto percebido no tempo	Reduziu um pouco	5	50,0
Impacto percebido no tempo	Não alterou	2	20,0
Impacto percebido no tempo	Aumentou um pouco	3	30,0
Intenção de continuidade	Definitivamente continuará	4	40,0
Intenção de continuidade	Provavelmente continuará	4	40,0
Intenção de continuidade	Ainda não sabe	1	10,0
Intenção de continuidade	Não pretende continuar	1	10,0

Fonte: Elaborado pelos autores (2026).

Nota: Rótulos abreviados. Alternativas sem marcação foram omitidas da tabela. Os percentuais foram arredondados para uma casa decimal. A expressão "diferença significativa" reproduz a alternativa do formulário e não indica teste de significância estatística.

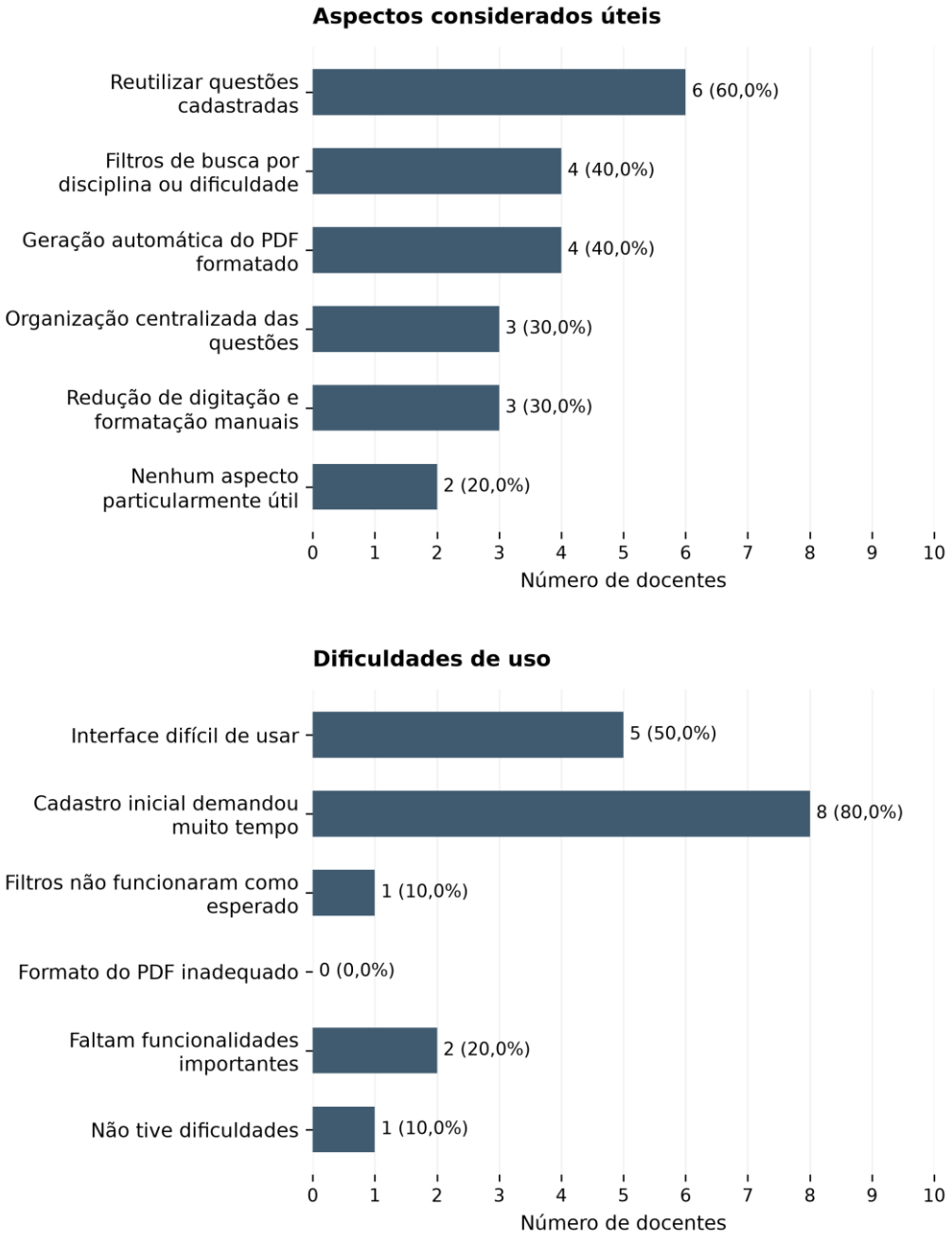
A possibilidade de reutilizar questões foi o aspecto útil mais assinalado, com seis marcações (60,0%), embora todos os respondentes tenham declarado questões novas na avaliação informada. O cadastro inicial foi a dificuldade mais frequente, com oito marcações (80,0%), seguido da interface, com cinco (50,0%). Dois docentes não identificaram aspecto particularmente útil (20,0%), e um declarou não ter dificuldades (10,0%). As demais marcações complementam a descrição dos aspectos úteis e das dificuldades (Figura 1).

A avaliação geral do módulo contou com nove respostas válidas: quatro participantes atribuíram nota 4 (44,4%), quatro nota 3 (44,4%) e um nota 5 (11,1%), em escala de 1, muito ruim, a 5, excelente. Não houve notas 1 ou 2. Uma resposta permaneceu ausente e não foi imputada.

Na grade de frequência, houve marcações para as cinco funcionalidades, mas as omissões variaram entre os recursos. O cadastro de questões não recebeu marcação em três registros; a montagem de provas, em cinco; e os filtros, a geração de PDF e a reutilização, em seis registros cada. Uma resposta indicou reutilização de ciclos anteriores na opção "Usei algumas vezes", em

contraste com a declaração de nenhuma questão reutilizada na avaliação informada. Também houve uma marcação de "Nunca usei" para a montagem de provas, apesar de o respondente informar uma avaliação elaborada com o módulo. A organização da grade e essas divergências limitaram a descrição da frequência de uso por funcionalidade.

Figura 1 – Aspectos considerados úteis e dificuldades de utilização do módulo (n = 10)



Fonte: Elaborado pelos autores (2026).

Nota: Perguntas de múltipla seleção; a soma dos percentuais pode ultrapassar 100% em cada conjunto.

4.4 Manifestações nas respostas abertas

Quatro docentes preencheram o comentário final. Dois campos complementares foram preenchidos por um desses mesmos participantes, totalizando seis manifestações textuais de quatro docentes. A leitura foi organizada nos eixos interface e fluxo de trabalho, esforço de cadastro e retrabalho, organização e institucionalização e condições de utilidade. Os eixos puderam coexistir numa mesma manifestação.

P02 destacou a organização da avaliação e a institucionalização de uma ferramenta para essa finalidade, mas também solicitou aperfeiçoamentos na interface e nas funcionalidades. P11 pediu melhor visualização das questões durante a montagem da prova. P13 apontou desafios na experiência de uso e sugeriu otimizar o fluxo de cadastro. Essas manifestações acompanharam as dificuldades de interface e cadastro assinaladas nas perguntas fechadas.

Uma manifestação crítica foi apresentada por P04, que descreveu a experiência como retrabalho para seu próprio processo. O mesmo participante informou que o processo ficou um pouco mais difícil e que não pretendia continuar utilizando o módulo. A resposta foi preservada em conjunto com as manifestações favoráveis e com a seleção de uma faixa de menor duração após o uso, em comparação com a habitual anterior.

5 DISCUSSÃO

5.1 Uso inicial e esforço de alimentação do banco

Os resultados situam a experiência investigada no momento inicial de utilização do módulo. Todos os dez docentes declararam cadastrar como novas as questões da avaliação informada, e oito apontaram o tempo demandado pelo cadastro inicial como dificuldade. Esse contexto é relevante para interpretar os relatos de melhoria, ausência de mudança e maior dificuldade, pois a experiência incluiu a alimentação do banco, além da montagem da prova.

A necessidade de trabalho para constituir um banco encontra correspondência na literatura, embora as atividades descritas não sejam idênticas. Beerepoot (2023) caracteriza como trabalhosa a construção de bancos compostos por itens de qualidade e destaca procedimentos de revisão e garantia da qualidade. Krzic e Brown (2022) discutem desafios relacionados à preparação de versões de questões, às orientações de correção e à seleção dos itens. Nos estudos citados, o esforço envolve a produção e a organização pedagógica do conteúdo; no PROINT, o questionário evidenciou especialmente a dificuldade percebida no cadastro inicial. A

aproximação entre essas experiências ajuda a compreender a constituição do banco como parte do trabalho docente, sem permitir estimar a equivalência do esforço envolvido.

A reutilização foi considerada útil por seis participantes, mas nenhuma questão foi declarada reutilizada na avaliação informada. Esse resultado pode ser interpretado como reconhecimento da possibilidade oferecida pela ferramenta. Não demonstra, contudo, economia de trabalho decorrente do reaproveitamento naquele ciclo. O acompanhamento de novas avaliações permitiria examinar se as questões cadastradas seriam efetivamente reutilizadas e como essa utilização se relacionaria ao esforço de preparação. A redução desse esforço em ciclos posteriores permanece uma hipótese a investigar.

A dificuldade relatada no cadastro inicial não deve ser interpretada apenas como resistência ao uso da ferramenta. Em experiências de constituição de bancos de questões, a produção, a revisão, a classificação e a organização dos itens demandam trabalho docente adicional antes que os benefícios da reutilização possam ser observados. No presente estudo, essa condição apareceu de forma explícita, pois todos os participantes declararam cadastrar questões novas na avaliação analisada.

5.2 Contribuições operacionais e alcance pedagógico

As marcações sobre geração automática de PDF, organização centralizada e redução de etapas manuais apontam contribuições percebidas para operações de preparação da prova. Esses resultados dialogam com um dos propósitos apresentados por Bearman, Nieminen e Ajjawi (2023), que tratam o digital como ferramenta para melhorar práticas de avaliação. No caso do PROINT, essa relação se manifesta no apoio à organização das questões e à composição do documento, conforme o recorte investigado.

O alcance dessas contribuições deve ser examinado em relação às decisões pedagógicas. No processo anterior, manter variedade e qualidade dos itens foi a dificuldade mais assinalada. A disponibilidade de filtros e de um documento formatado pode apoiar o trabalho, mas o instrumento aplicado não examinou se as questões se tornaram mais claras, mais adequadas aos objetivos de aprendizagem ou mais consistentes quanto à dificuldade. A ênfase de Beerepoot (2023) na revisão e na garantia da qualidade dos itens reforça a pertinência de considerar essas dimensões em avaliações futuras do banco.

Também é necessário distinguir as funções dos sistemas comparados. Parte das vantagens discutidas por Krzic e Brown (2022) refere-se a avaliações online com correção

automática. O recorte do PROINT focaliza o cadastro, a seleção e a montagem da prova em PDF. Assim, benefícios relativos à correção automatizada não podem ser transferidos para explicar os resultados deste estudo. A contribuição aqui examinada diz respeito ao processo de preparação relatado pelos docentes.

5.3 Diversidade das experiências e interpretação do tempo

As respostas não indicam uma experiência uniforme. Quatro docentes perceberam melhora no processo, três não perceberam diferença significativa e três relataram maior dificuldade. A intenção de continuidade de oito participantes coexistiu com problemas de interface e com o esforço de cadastro. Essa combinação sugere que reconhecer utilidade na ferramenta e encontrar dificuldades durante o uso são experiências que podem ocorrer simultaneamente.

Barbosa, Silva e Freitas (2022) também recorreram a questionário para examinar a adequação de um banco de questões aos seus objetivos e a intenção de utilização. Essa aproximação metodológica evidencia o interesse das manifestações dos docentes na avaliação de ferramentas desse tipo. No PROINT, entretanto, a intenção de continuidade expressa uma disposição declarada, que precisa ser diferenciada da utilização efetiva em semestres posteriores e da demonstração de melhoria de desempenho.

As divergências entre faixas de tempo e impacto percebido constituem um resultado importante da comparação individual. Uma faixa posterior de menor duração não foi acompanhada, em todos os casos, de percepção de redução. Da mesma forma, houve relato de pequena redução dentro da mesma faixa. Os dois itens possuem referências distintas: a rotina habitual anterior e a avaliação elaborada com o módulo. Além disso, as categorias não registram variações internas aos intervalos, e o número e a complexidade das questões das avaliações anteriores não foram controlados. O desenho não permite determinar quais desses fatores explicam cada divergência.

Os cinco relatos de pequena redução de tempo expressam a percepção dos participantes e devem ser lidos em conjunto com essas diferenças. A manutenção das respostas discordantes permite descrever a experiência sem selecionar apenas o indicador mais favorável à ferramenta.

5.4 Implicações para o artefato e limitações do estudo

Na perspectiva da DSR, a avaliação deve relacionar os objetivos da solução às evidências obtidas em sua utilização. Peffers et al. (2007) incluem manifestações dos usuários e pesquisas de satisfação entre as possibilidades de avaliação e admitem que os resultados orientem aperfeiçoamentos do artefato. Nesse sentido, as dificuldades de cadastro, as solicitações de melhor visualização das questões e os comentários sobre a interface oferecem subsídios para priorizar revisões no módulo. O estudo não examinou a implementação dessas revisões nem seus efeitos posteriores.

A análise também sugere que a avaliação institucional do módulo pode abranger dimensões complementares: esforço de cadastro, facilidade de localização e seleção dos itens, montagem do documento e qualidade pedagógica das questões. As primeiras dimensões se aproximam das experiências relatadas no questionário; a última demanda procedimentos específicos que não foram realizados nesta investigação. A separação dessas dimensões pode tornar mais precisos os objetivos de futuras avaliações da ferramenta.

Os resultados estão circunscritos a uma instituição e a um grupo de adesão voluntária, com dez participantes na comparação entre os momentos. Os dois docentes incluídos no bloco A sem resposta posterior não podem ser caracterizados como satisfeitos ou insatisfeitos com o módulo. A ausência de grupo de controle, o autorrelato, as diferentes faixas de tempo e a falta de controle da extensão e da complexidade das provas limitam a atribuição das mudanças à ferramenta. As omissões e a organização da grade de frequência restringem, ainda, a descrição de utilização por funcionalidade.

A participação do autor no desenvolvimento do módulo constitui outra condição relevante, por sua possível influência na relação com os docentes e na interpretação dos dados. A preservação das manifestações críticas e o registro das decisões de tratamento favorecem a transparência, sem eliminar essa possibilidade. O registro das diferentes experiências delimita a contribuição do estudo à utilização inicial do artefato nesse contexto institucional.

6 CONSIDERAÇÕES FINAIS

A principal contribuição deste estudo não consiste em demonstrar uma redução mensurada do tempo de elaboração das avaliações, mas em apresentar evidências contextualizadas sobre a experiência inicial de docentes com um banco de questões institucional. Os resultados mostram que a ferramenta foi associada a contribuições

operacionais percebidas, especialmente na organização das questões, na montagem das provas e na geração de documentos em PDF, ao mesmo tempo em que evidenciaram obstáculos relacionados ao cadastro inicial e à interface.

Sendo assim, este estudo analisou em que medida a utilização do banco de questões integrado ao PROINT influenciou o processo de elaboração de avaliações pelos docentes do Centro Universitário Santa Terezinha. No contexto investigado, os relatos indicaram contribuições percebidas para a organização das questões, a montagem das provas e a geração do documento em PDF, acompanhadas de dificuldades relacionadas ao cadastro inicial e à interface. A influência relatada variou entre os participantes, incluindo experiências de melhoria, ausência de mudança e maior dificuldade.

A percepção de pequena redução de tempo coexistiu com divergências entre as faixas informadas e a avaliação direta desse impacto. Os dados permitiram descrever mudanças percebidas, sem quantificar uma economia de minutos atribuível ao módulo. A intenção de continuidade predominou, acompanhada de manifestações críticas e demandas de aperfeiçoamento.

O uso inicial constitui uma condição central de interpretação, pois todos os respondentes do segundo momento declararam questões novas na avaliação informada. Os resultados abrangem, assim, a alimentação do banco e a preparação da prova. O reconhecimento da reutilização como aspecto útil não demonstra benefícios acumulados em ciclos posteriores.

A contribuição do trabalho consiste em apresentar evidências contextualizadas sobre a experiência docente com um artefato institucional e em identificar demandas para seu aprimoramento, especialmente no fluxo de cadastro e na visualização das questões durante a montagem das avaliações. Essas evidências podem subsidiar decisões institucionais, mas não demonstram melhoria da qualidade pedagógica dos itens ou da aprendizagem dos estudantes, dimensões que não foram avaliadas.

O grupo de dez participantes pareados, a adesão voluntária, o autorrelato e a realização em uma única instituição delimitam o alcance das conclusões. As restrições do instrumento e a participação do autor no desenvolvimento do módulo, discutidas anteriormente, também devem ser consideradas na leitura dos resultados.

Como continuidade, propõe-se acompanhar os mesmos docentes em novos ciclos avaliativos, distinguindo o esforço de cadastro inicial do trabalho de revisão e reutilização das questões. A adoção de categorias de tempo equivalentes, o registro mais preciso das

características das provas e a combinação dos questionários com observação do processo ou registros de uso podem ampliar a compreensão das contribuições e das dificuldades do módulo. Estudos posteriores também poderão examinar a qualidade dos itens e avaliar os efeitos de eventuais aperfeiçoamentos da ferramenta.

REFERÊNCIAS

BARBOSA, Pedro Luis Saraiva; SILVA, Roberta da; FREITAS, Mateus de Lima. Banco de questões do Instituto Federal do Ceará: um sistema web para auxiliar o processo de avaliação do estudante no ensino superior. **Revista Internacional de Educação Superior**, Campinas, v. 8, e022006, 2022. https://doi.org/10.20396/riesup.v8i00.8657696.

BEARMAN, Margaret; NIEMINEN, Juuso Henrik; AJJAWI, Rola. Designing assessment in a digital world: an organising framework. **Assessment & Evaluation in Higher Education**, v. 48, n. 3, p. 291–304, 2023. https://doi.org/10.1080/02602938.2022.2069674.

BEEREPOOT, Maarten T. P. Formative and summative automated assessment with multiple-choice question banks. **Journal of Chemical Education**, v. 100, n. 8, p. 2947–2955, 2023. https://doi.org/10.1021/acs.jchemed.3c00120.

GALVÃO, Nadielli Maria dos Santos; MADUREIRA, Jamille Silva; SCHNEIDER, Henrique Nou. Design Science Research para o desenvolvimento de artefatos educacionais. **Boletim de Conjuntura (BOCA)**, Boa Vista, v. 18, n. 54, p. 194–220, 2024. https://doi.org/10.5281/zenodo.13119424.

GONZAGA, Antônia Edivaneide de Sousa et al. A avaliação da aprendizagem na formação docente: revelando horizontes. **Estudos em Avaliação Educacional**, São Paulo, v. 34, e08097, 2023. https://doi.org/10.18222/eae.v34.8097.

KENNY, John; FLUCK, Andrew Edward. Life at the academic coalface: validation of a holistic academic workload estimation tool. **Higher Education**, v. 86, p. 507–526, 2023. https://doi.org/10.1007/s10734-022-00912-x.

KRZIC, Maja; BROWN, Sandra. Question banks for effective online assessments in introductory science courses. **Natural Sciences Education**, v. 51, n. 2, e20091, 2022. https://doi.org/10.1002/nse2.20091.

NAMESNIK-SILVESTER, Kirsten et al. ShareStats: an open statistics item bank developed by a community of instructors in higher education. **Teaching Statistics**, v. 47, n. 2, p. 129–138, 2025. https://doi.org/10.1111/test.12399.

PEFFERS, Ken et al. A design science research methodology for information systems research. **Journal of Management Information Systems**, v. 24, n. 3, p. 45–77, 2007. https://doi.org/10.2753/MIS0742-1222240302.

PEZZI, Marcelo Rodrigo; VEIGA, Amélia; FERNANDES, Preciosa. Práticas avaliativas dos professores do ensino superior: uma análise crítica de seus impactos na aprendizagem. **Revista FAFIRE**, Recife, v. 18, n. 1, p. 125–141, 2025. https://doi.org/10.24024/23585188v18n1a2025p01250141.

VIBERG, Olga et al. Exploring teachers' (future) digital assessment practices in higher education: instrument and model development. **British Journal of Educational Technology**, v. 55, n. 6, p. 2597–2616, 2024. https://doi.org/10.1111/bjet.13462.