

<https://doi.org/10.18222/eae.v34.9951>

RELAÇÃO ENTRE A COMPETÊNCIA E A EXPERIÊNCIA EM AVALIAÇÃO DOS AVALIADORES DO SINAES

 LUIS CARLOS DA SILVA CASTRO^I

 ROBERTO BRAZILEIRO PAIXÃO^{II}

^I Universidade Federal da Bahia (UFBA), Salvador-BA, Brasil; castro.luis@ufba.br

^{II} Universidade Federal da Bahia (UFBA), Salvador-BA, Brasil; roberto.brazileiro@ufba.br

RESUMO

Este estudo busca analisar a relação entre a competência e a experiência em avaliação dos avaliadores do Sistema Nacional de Avaliação da Educação Superior (Sinaes). Foi feita uma pesquisa Delphi para listar as competências em avaliação consideradas mais adequadas ao contexto do Sinaes. Na sequência foi realizada uma *survey* para identificar a experiência em avaliação e a autoavaliação deles sobre o desenvolvimento de suas competências. Os dados empíricos evidenciaram que, em todas as dimensões teóricas de competências, quanto maior a quantidade de avaliações feitas, maior o nível de desenvolvimento dessas competências. Recomenda-se que avaliadores mais experientes trabalhem em parceria com outros menos experientes, a fim de desenvolver as competências em avaliação.

PALAVRAS-CHAVE AVALIAÇÃO • COMPETÊNCIAS • EXPERIÊNCIA • SINAES.

COMO CITAR:

Castro, L. C. da S., & Paixão, R. B. (2023). Relação entre a competência e a experiência em avaliação dos avaliadores do Sinaes. *Estudos em Avaliação Educacional*, 34, Artigo e09951, 2023. <https://doi.org/10.18222/eae.v34.9951>

RELACIÓN ENTRE COMPETENCIA Y EXPERIENCIA SEGÚN LA EVALUACIÓN DE LOS EVALUADORES DEL SINAES

RESUMEN

Este estudio intenta analizar la relación entre la competencia y la experiencia en evaluación de los evaluadores do Sistema Nacional de Avaliação da Educação Superior [Sistema Nacional de Evaluación de la Educación Superior] (Sinaes). Se efectuó una pesquisa Delphi para presentar las competencias en evaluación consideradas más adecuadas al contexto del Sinaes. Después de ello se llevó a cabo una *survey* para identificar su experiencia en evaluación y su autoevaluación sobre el desarrollo de sus competencias. Los datos empíricos pusieron de manifiesto que, en todas las dimensiones teóricas de competencias, cuanto mayor es la cantidad de evaluaciones realizadas, mayor el nivel de desarrollo de dichas competencias. Se recomienda que los evaluadores más experimentados trabajen junto con otros menos experimentados, con el propósito de desarrollar las competencias en evaluación.

PALABRAS CLAVE EVALUACIÓN • COMPETENCIAS • EXPERIENCIA • SINAES.

RELATIONSHIP BETWEEN SINAES EVALUATORS' COMPETENCE AND EXPERIENCE IN EVALUATION

ABSTRACT

This study seeks to analyze the relationship between the competence and experience in evaluation of the Sistema Nacional de Avaliação da Educação Superior [National System for the Evaluation of Higher Education evaluators] (Sinaes). A Delphi study was carried out to list the evaluation competencies considered most appropriate to the context of the Sinaes. A survey was then carried out to identify their experience in evaluation and their self-assessment of the development of their competences. The empirical data showed that in all the theoretical dimensions of competences the greater the number of assessments carried out, the greater the level of development of these competences. It is recommended that more experienced evaluators work in partnership with less experienced ones in order to develop evaluation competencies.

KEYWORDS EVALUATION • COMPETENCIES • EXPERIENCE • SINAES.

Recebido em: 8 DEZEMBRO 2022

Aprovado para publicação em: 12 JULHO 2023



Este é um artigo de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY-NC.

INTRODUÇÃO

O processo de avaliação externa de uma instituição de ensino superior brasileira, no âmbito do Sistema Nacional de Avaliação da Educação Superior (Sinaes), apresenta especificidades que demandam dos avaliadores, naturalmente, o desenvolvimento de competências em avaliação. O Sinaes dispõe de um Banco de Avaliadores (BASis – Banco de Avaliadores do Sistema Nacional de Avaliação da Educação Superior) formado por docentes da educação superior com vínculo institucional público ou privado. Eles são designados para avaliar a qualidade das instituições e de seus cursos de graduação, conforme a Portaria Normativa n. 840, de 24 de agosto de 2018 do Ministério da Educação (Portaria Normativa n. 840, 2018).

Ao realizar as avaliações nas Instituições de Ensino Superior (IES) do país, os avaliadores do Sinaes precisam estar cientes da importância do pleno desenvolvimento de suas competências em avaliação, para propiciar avaliações eficazes, o que implica benefícios para a própria instituição e para toda a sociedade. Além disso, sua experiência na execução de avaliações contribui para que esse desenvolvimento aconteça de maneira eficiente. Dessa forma, esta pesquisa¹ tem como objetivo analisar a relação entre a experiência e as competências em avaliação dos avaliadores pertencentes ao BASis.

Entender como o pleno desenvolvimento de competências em avaliação se relaciona com a experiência em avaliação pode ajudar a definir melhores políticas de designação de avaliadores, além de elevar o nível de competências do próprio BASis, ou seja, na ampliação da capacidade avaliativa (*capacity building*) do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Um exemplo é a designação de avaliadores com maior e menor experiência a fim de elevar as competências em avaliação dos integrantes recentes do BASis. Considerando que esse banco é dinâmico, entende-se que apenas a capacitação formal em avaliação pode não ser suficiente para a profissionalização do avaliador, sendo a inserção prática fundamental. Logo, pressupõe-se que novos avaliadores do sistema poderiam aperfeiçoar suas habilidades ao atuar com avaliadores mais experientes.

COMPETÊNCIAS E EXPERIÊNCIA EM AVALIAÇÃO

A literatura de competências em avaliação descreve que existe uma relação direta entre a experiência em avaliação do avaliador e o pleno desenvolvimento das competências na prática da avaliação (Kaesbauer, 2012). Embora o conhecimento básico para o desenvolvimento dessas habilidades seja um curso de avaliação tradicional, a aplicação do conhecimento adquirido no contexto de uma avaliação real é fundamental (Trevisan, 2004). Um avaliador pode ter todos os conhecimentos necessários

1 Este artigo é derivado da dissertação de mestrado de Castro (2021).

para conduzir uma avaliação e ainda assim não atuar de forma satisfatória, uma vez que as habilidades adquiridas na capacitação precisam ser refinadas por meio da prática (European Evaluation Society [EES], 2011).

As experiências adquiridas por meio de atividades desenvolvidas no contexto real de avaliação – como trabalhar em projetos de avaliação, desenvolver e divulgar protocolos de avaliação, realizar entrevistas, conduzir análise de dados e redigir relatórios de avaliação – se traduzem numa oportunidade para que os concluintes de um curso de avaliação entendam a força e o alcance que a experiência em avaliação pode ter em sua formação como avaliador (Trevisan, 2004).

As listas de competências em avaliação que serviram de base para a realização deste estudo foram elaboradas pelos acadêmicos King et al. (2001), pela American Evaluation Society (AEA, 2018), Canadian Evaluation Society (CES, 2018), Australasian Evaluation Society (AES, 2013), EES (2011), UK Evaluation Society (UKES, 2012) e Aotearoa New Zealand Evaluation Association (Anzea, 2011), os quais abordam um rol de competências em avaliação e classificam-nas em dimensões teóricas específicas de acordo com seus aspectos conceituais e também profissionais, dos quais experiência faz parte.

O primeiro grande estudo a abordar os aspectos teóricos e empíricos de competências em avaliação foi a lista de Competências Essenciais para Avaliadores de Programas elaborada a partir do trabalho de King et al. (2001). Essas competências estão categorizadas em seis domínios de prática ou dimensões teóricas: Prática Profissional (competências relacionadas a normas e valores fundamentais de uma avaliação); Investigação Sistemática (relacionadas aos aspectos técnicos da avaliação); Análise Situacional (relacionadas à análise e ao atendimento de interesses, questões e circunstâncias contextuais únicas de uma avaliação); Gerenciamento de Projetos (relacionadas às ferramentas usadas na avaliação); Prática Reflexiva (relacionadas à consciência da própria experiência em avaliação); e Competência Interpessoal (relacionadas às habilidades pessoais utilizadas na realização das avaliações); sendo que cada uma está relacionada a um conjunto de aspectos necessários para a prática eficaz da avaliação (Stevahn et al., 2005a).

Associações e sociedades de avaliação

Os estudos de King et al. (2001) impulsionaram o surgimento de outras listas de competências, cujas dimensões teóricas refletiram os aspectos profissionais e o contexto do próprio processo de avaliação elencados por cada associação ou sociedade de avaliação.

A lista de competências da Anzea foi publicada em setembro de 2011 e é composta de quatro domínios: Análise Contextual e Engajamento, na qual são descritas as habilidades críticas para a análise do contexto do processo avaliativo e o

envolvimento das partes interessadas na avaliação; Inquérito Avaliativo Sistemático, no qual se pensa cuidadosa e criticamente sobre o que, por que e como se faz a avaliação; Gerenciamento de Projetos de Avaliação e Prática de Avaliação Profissional, em que são descritas as competências necessárias para gerenciar uma avaliação de maneira profissional; e Prática Reflexiva e Desenvolvimento Profissional, que inclui competências que contribuem para o desenvolvimento do profissional e da profissão de avaliação. Vale ressaltar que o fato de as competências da Anzea são diferentes daquelas elaboradas por outras associações de avaliação, por serem focadas nos aspectos relacionados ao processo da avaliação e de ser rigorosa quanto à incorporação da cultura e valores elencados no Tratado de Waitangi/Direitos Indígenas as tornam diferentes daquelas elaboradas por outras associações de avaliação (Wehipeihana et al., 2014).

A lista de competências da EES, publicada em 2011, teve como principal objetivo fortalecer uma cultura de profissionalismo e estimular a boa prática da avaliação, levando em consideração fatores como a consciência de profissionais e clientes, a dimensão do interesse público da avaliação, o fortalecimento do senso de identidade entre os avaliadores e a responsabilidade dos avaliadores no sentido de agregar valor a seus clientes e à sociedade (EES, 2011). A lista publicada pela EES é composta de três grupos de domínios da avaliação: Conhecimento de Avaliação, que elenca competências relacionadas à compreensão adequada do histórico da avaliação, abordagens, modelos, teorias e suas implicações em relação à governança da avaliação, *design*, propósitos, práticas e métodos e os diversos usos da avaliação na sociedade; Prática Profissional, que abrange competências relacionadas ao que realmente acontece no campo; e Disposições e Atitudes de Avaliação, que contém competências relacionadas às qualidades pessoais que permitem que os avaliadores pratiquem a disciplina não apenas em circunstâncias receptivas, mas também em contextos difíceis (EES, 2011).

As competências elencadas pela Sociedade de Avaliação do Reino Unido (UKES), publicadas em junho de 2012, têm como principal objetivo inserir os padrões e códigos de conduta do avaliador em na sua prática profissional. A estrutura observada distribui competências gerais em avaliação de acordo com três categorias principais: Conhecimento de Avaliação, que se refere tanto à base de conhecimento da filosofia quanto ao método que a avaliação traz das ciências sociais, disciplinas e profissões e de conhecimentos específicos gerados no campo da avaliação; Prática Profissional, que se refere às competências no campo, onde as capacidades do avaliador precisam garantir que ele possa não só conduzir uma avaliação confiável e válida, mas também que tenha as habilidades interpessoais e políticas para gerenciar o processo da avaliação; e Qualidades e Atitudes, referente às características pessoais que permitem que os avaliadores atuem em circunstâncias difíceis, que

mantenham uma postura imparcial e independente, gerenciem a complexidade e tolerem ambiguidades e dissonâncias de um processo avaliativo (UKES, 2012).

A lista de competências publicada pela Sociedade de Avaliação Australiana (AES) em 2012 tem o objetivo de orientar e apoiar os membros e outras partes interessadas a otimizarem seus conhecimentos e experiência em avaliação (AES, 2013). As principais categorias dessa lista de competências são: Atitude Avaliativa e Prática Profissional, que abrange competências focadas no desenvolvimento profissional contínuo e de atitudes avaliativas pessoais e grupais; Fundamentos Teóricos, que se refere aos aspectos teóricos da avaliação; Atenção à Cultura, Partes Interessadas e Contexto, que contempla as competências relacionadas à multiplicidade de perspectivas de valor, incluindo valores culturais, sociais e políticos que cercam o avaliador; Métodos de Pesquisa e Investigação Sistemática, que elenca competências que levam em consideração a necessidade do conhecimento e as habilidades em métodos de pesquisa e investigação sistemática para a coleta de dados válidos e confiáveis; Gerenciamento de Projetos de Avaliação, associada às habilidades de gerenciamento de projetos das quais os avaliadores precisam para negociar, definir o escopo, gerenciar e concluir uma avaliação efetivamente; Habilidades Interpessoais, que diz respeito às habilidades para se comunicar efetivamente com clientes, consumidores e outras partes interessadas em uma avaliação; e Atividades de Avaliação, que se concentram nas tarefas que um avaliador executará no decorrer de uma avaliação (AES, 2013).

A relação de competências da AEA, publicada em 2018, tem o objetivo de aprimorar práticas e métodos de avaliação, aumentando seu uso e promovendo-a como profissão, apoiando também a geração eficaz de teoria e conhecimento sobre a ação humana. Essa abordagem reflete a preocupação dos membros da AEA com a autorreflexão sobre pontos fortes, limitações da avaliação e a busca de maneiras apropriadas de expandir e melhorar a avaliação (AEA, 2018). Os domínios da AEA são: Prática Profissional, que se concentra no que torna os avaliadores distintos como praticantes profissionais; Metodologia, que tem foco nos aspectos técnicos da investigação sistemática da avaliação, a qual é baseada em evidências; Contexto, que compreende as circunstâncias únicas, múltiplas perspectivas e as mudanças nas configurações das avaliações e seus usuários/partes interessadas; Planejamento e Gestão, que compreende as competências que permitem ao avaliador determinar e monitorar planos de trabalho, cronogramas, recursos e outros componentes necessários para concluir e entregar um estudo de avaliação; e Interpessoal, que tem foco nas relações humanas e interações sociais que fundamentam a eficácia do avaliador para a prática profissional ao longo da avaliação (AEA, 2018).

A experiência profissional no campo da avaliação é um dos indicadores que contribui para categorizar os avaliadores de programas de acordo com seu nível de

expertise em relação ao desenvolvimento de competências em avaliação. Por isso, as associações de avaliação e os autores das principais listas de competências não só levam em consideração, nos seus estudos, o conhecimento e a experiência, mas também reconhecem que as competências podem ser direcionadas a diferentes estágios de desenvolvimento, desde o nível básico, requisito de entrada na atividade de avaliação, até aqueles especializados em avaliação (EES, 2011).

De acordo com Stevahn et al. (2005a), um avaliador iniciante tem experiência limitada, o proficiente tem experiência moderada e aquele considerado *expert* tem grande experiência profissional em avaliação, tornando possível afirmar que a experiência profissional relaciona-se diretamente com o desenvolvimento de competências em avaliação evidenciado por um avaliador.

King e Stevahn (2015) sinalizam a existência de poucas pesquisas que associam competências específicas às práticas de avaliação. Para esses autores, a realização de pesquisas que abordem as competências do avaliador representaria uma oportunidade para a descoberta de formas úteis de ensinar aos avaliadores iniciantes as competências em avaliação necessárias para a formação de boas práticas de avaliação no decorrer do tempo (Stevahn et al., 2005b).

MÉTODO

A operacionalização da pesquisa foi conduzida em duas fases. Primeiro, foi realizada uma pesquisa Delphi com 101 competências obtidas dos trabalhos de King et al. (2001), da American Evaluation Society (2018), da Canadian Evaluation Society (2018), da Australasian Evaluation Society (2013), da Aotearoa New Zealand Evaluation Association (2011), da European Evaluation Society (2011) e da UK Evaluation Society (2012), com o objetivo de elencar aquelas consideradas mais adequadas ao contexto do Sinaes. Essa listagem foi elaborada com base na análise de todas as competências em avaliação elencadas nas listas abordadas nesta pesquisa, com alguns agrupamentos feitos de acordo com a semelhança dos conceitos e objetivos no campo da avaliação e posterior classificação nas dimensões teóricas da lista de Competências Essenciais para Avaliadores de Programas de King et al. (2001).

Vale ressaltar que, embora as listas de competências da EES e da UKES apresentem dimensões teóricas diferentes das outras listas elencadas nesta pesquisa, a natureza das competências em avaliação pertencentes a elas também possibilitou seu enquadramento nas dimensões da lista de King et al. (2001), a qual se tornou parâmetro para as listas de competências em avaliação abordadas neste estudo.

Convém destacar que a pesquisa Delphi foi criada em 1962 por Dalkey e Helmer e constitui um método para obter o consenso de ideias de um grupo de especialistas sobre determinado objeto de estudo, através de uma série de questionários inter-

calados com *feedback* controlado das opiniões (Dalkey & Helmer, 1962). Vale ressaltar que a seleção de sujeitos (especialistas), prazos para a realização e conclusão do estudo, possibilidade de baixa taxa de respostas e *feedback* não intencional do grupo respondente são pontos de atenção no método (Hsu & Sandford, 2007).

Na Fase 2, foi conduzida uma *survey* com o público-alvo, com o objetivo de identificar sua experiência em avaliação e sua autoavaliação acerca das competências em avaliação consideradas já mais ou menos desenvolvidas.

O público-alvo foi formado pelos avaliadores do BASis aprovados no curso de capacitação e convocados mediante portarias de ingresso emitidas pela Diretoria de Avaliação da Educação Superior (Daes) do Inep nos anos de 2018 e 2019. Eles representam os avaliadores mais recentemente habilitados a realizar as avaliações do Inep e totalizaram 5.951 docentes. No total, o BASis conta com 10.562 avaliadores (Diretoria de Avaliação da Educação Superior [Daes], 2021).

Procedimentos metodológicos da Fase 1

Na realização da Delphi, as competências em avaliação foram organizadas em um questionário e apresentadas por *e-mail* a cinco especialistas, os quais constavam na lista de avaliadores do BASis nos anos de 2018 e 2019. Essa quantidade foi definida por ser um número ímpar, para que o dissenso fosse evitado, já que o objetivo foi garantir uma opinião majoritária dos participantes acerca das competências consideradas mais adequadas ao contexto de avaliação do Sinaes.

Não há na literatura um padrão estabelecido para os critérios de escolha dos participantes de uma Delphi (Hsu & Sandford, 2007). Nesse sentido, os especialistas selecionados foram profissionais com conhecimento e experiência relevantes na execução das avaliações do Sinaes. Os especialistas foram selecionados por conveniência a partir da rede de contatos dos autores, tendo como critérios ter doutorado completo e já ter participado de mais de dez avaliações do Sinaes. Vale ressaltar que o processo de execução da Delphi foi adaptado do estudo de Diaz et al. (2020) e aconteceu em três rodadas, todas elas realizadas de forma sequencial.

- a) Primeira rodada – foi iniciada com a coleta da idade e número de avaliações do Sinaes realizadas pelo especialista e continuou com a realização da seguinte pergunta: “As competências em avaliação listadas abaixo são fundamentais para o processo avaliativo do Sinaes? Marque um ‘x’ naquelas que você considera que são fundamentais para o trabalho do avaliador do S”. Foram levadas para a segunda rodada apenas as competências marcadas com o “x” pelos especialistas.
- b) Segunda rodada – foi apresentado aos especialistas um relatório com a frequência relativa de respostas nas competências selecionadas na 1ª rodada e os comentários feitos na rodada anterior. Em seguida, foi feita a seguinte

pergunta: “Qual grau de importância você atribuiria a cada uma das competências em avaliação relacionadas abaixo para um bom desempenho dos avaliadores na execução das avaliações do Sinaes?”. A escala teve como variação: 1 – Muito importante; 2 – Importante; 3 – Pouco importante; e 4 – Nada importante. As competências que foram levadas para a rodada subsequente foram as que dois terços ou mais dos especialistas atribuíram o grau 1 ou 2. Vale ressaltar que, em virtude da desistência de um dos especialistas nesta rodada, houve o convite, e posterior aceite, de outro. Optou-se por convidar um profissional com doutorado e experiência de mais de 25 anos na gestão de Instituições de Ensino Superior, tendo sido reitor de mais de uma e, consequentemente, com histórico de recepção de avaliações do Ministério da Educação (MEC). A ideia foi incorporar a visão não apenas do avaliador, mas também de quem é responsável pela instituição avaliada.

- c) Terceira rodada – os especialistas tiveram acesso, de maneira prévia, ao relatório da frequência relativa de respostas e comentários feitos em cada competência selecionada na segunda rodada. Para efeito de comparação com o resultado global de cada competência, a resposta assinalada por cada especialista em cada uma delas na rodada anterior foi sinalizada de azul. A seguinte pergunta foi feita nesta rodada: “Você concorda que cada uma das competências em avaliação listadas abaixo é importante para um bom desempenho dos avaliadores na execução das avaliações do Sinaes?” A escala de resposta teve como variação: 1 – Concordo plenamente; 2 – Concordo; 3 – Concordo parcialmente; 4 – Discordo; 5 – Discordo parcialmente; e 6 – Discordo totalmente. Dentro de cada resposta, foram consideradas consenso as competências com as quais dois terços dos especialistas concordaram plenamente ou apenas concordaram acerca da sua adequação ao contexto do Sinaes.

Com o objetivo de possibilitar um melhor entendimento dos critérios adotados para a seleção de competências na segunda e na terceira rodadas, ressalta-se que foi utilizada a regra matemática do arredondamento de números decimais. Utilizou-se esse recurso para otimizar a seleção das competências na segunda e na terceira rodadas, uma vez que a aplicação de dois terços no total de respostas nessas rodadas, que foi igual a cinco (quantidade de especialistas participantes da Delphi), tem como resultado 3,3 (três inteiros e três décimos). Esse número, de acordo com a regra supracitada, foi arredondado para três inteiros. Nesse sentido, a competência cuja soma de respostas nos níveis predefinidos para a seleção teve como resultado três avançou para a etapa subsequente. Em relação à variação da escala utilizada na terceira rodada, foi retirado o nível “Não concordo nem discordo”, já que, de acordo com Halaydna e Rodriguez (2013), a probabilidade do alcance

de um resultado mais eficaz na seleção das competências aumenta quando se eliminam níveis intermediários.

Dos cinco participantes da Delphi, apenas quatro responderam ao questionário enviado na terceira e última rodada. Em virtude da inércia do especialista não respondente em relação às tentativas de contato, a análise dos resultados ocorreu, apenas, com as respostas de quatro especialistas. Os procedimentos elencados a seguir foram adotados como uma tentativa de minimizar os efeitos que a ausência de respostas de um dos respondentes poderia provocar na seleção das competências em avaliação:

- a) os critérios iniciais definidos na terceira rodada foram aplicados ao resultado consolidado das respostas obtidas pelos quatro participantes, e as competências que atendiam a esses critérios foram selecionadas;
- b) das competências não selecionadas, aquelas que não atenderiam aos critérios iniciais de seleção definidos nessa rodada, mesmo que o participante não respondente continuasse na Delphi e marcasse qualquer uma das categorias de concordância, foram descartadas;
- c) as competências que restaram após a realização dos procedimentos “a” e “b” foram aquelas cuja probabilidade de serem selecionadas eram reais, caso o especialista não respondente continuasse na Delphi. Dessa lista de competências restantes, apenas aquelas cuja soma das respostas nas categorias “Concordo plenamente”, “Concordo” e “Concordo parcialmente” foi maior que nas categorias “Discordo parcialmente”, “Discordo” e “Discordo totalmente” foram selecionadas. A aplicação desse critério possibilitou, somente, a seleção daquelas competências nas quais houve mais consenso do que divergências entre os especialistas.

A lista de 63 competências em avaliação que restou, após a realização dos procedimentos descritos nos itens “a”, “b” e “c”, na terceira rodada da Delphi, foi aquela utilizada na pesquisa de campo com os avaliadores externos do Sinaes.

Procedimentos metodológicos da Fase 2

A coleta de dados junto aos avaliadores foi feita por meio de uma pesquisa *on-line* com uso da plataforma SurveyMonkey. No questionário, as competências em avaliação validadas pela Delphi foram distribuídas por dimensão teórica, em que as respostas foram categorizadas em: 1 – Discordo totalmente; 2 – Discordo; 3 – Concordo; e 4 – Concordo totalmente. O respondente assinalava em que medida ele tinha plenamente desenvolvida cada competência elencada. Optou-se pela escala de quatro pontos para evitar posições intermediárias (Halaydna & Rodriguez, 2013).

A frequência de respostas em determinada classificação da escala evidenciou o grau de concordância em relação ao pleno desenvolvimento de competências em avaliação autodeclarado pelo avaliador. Vale ressaltar que, embora possibilite uma dispersão menor de respostas, a opção por utilizar uma escala com número reduzido de categorias se deu em função da quantidade de itens levados a campo, uma vez que questionários com número de itens e variações de escala relativamente altos podem gerar baixa taxa de respostas (Halaydna & Rodriguez, 2013). Foram também coletados dados relacionados a idade, gênero e nível de escolaridade dos avaliadores externos.

O questionário foi previamente aplicado a um grupo de cinco profissionais da área da educação superior, entre os quais dois docentes, para a verificação da necessidade de possíveis ajustes antes de sua aplicação aos avaliadores externos. Após a análise dos participantes desse processo de validação de conteúdo, entre outras sugestões apontadas de natureza gramatical, foi verificado que o significado da competência “Praticar dentro da própria competência”, da categoria Prática Profissional, não ficou claro para dois deles. A partir desse *feedback*, a referida competência foi retirada do questionário, a fim de minimizar o efeito de uma possibilidade relativamente alta de não respostas nesse item, em específico, decorrente do não entendimento do seu conceito. Além disso, adaptações de natureza gramatical foram feitas em algumas competências elencadas.

Aspectos relacionados à amostra

O critério utilizado para o levantamento dos *e-mails* dos avaliadores inclusos na população da pesquisa – para o envio do questionário e a posterior obtenção da amostra a ser utilizada para análise dos dados empíricos – foi o do acesso aos dados disponíveis em suas produções acadêmicas e nas instituições de ensino onde lecionam, o que foi feito mediante consulta inicial ao currículo Lattes de cada um. A quantidade de docentes avaliadores que representou a população da pesquisa (5.951) foi dividida entre cinco assistentes de pesquisa, os quais, de forma preliminar, buscaram informações de natureza acadêmica e profissional de todos eles, e assim obtiveram a maior quantidade de *e-mails* possível, no período de janeiro a março de 2021, o que resultou em 4.130 *e-mails* coletados.

O questionário foi, então, enviado aos 4.130 docentes, sendo as respostas coletadas entre junho e julho de 2021. Ao todo, foram registradas 974 respostas. Desse total, foram excluídos participantes que não aceitaram responder, não finalizaram ou não responderam nas escalas de concordância, bem como aqueles cujas respostas foram dadas em sequência (todas iguais). Após a exclusão dessas respostas, obteve-se uma amostra final de 529 respondentes. Vale ressaltar que, por ter sido uma amostra obtida pelo critério da conveniência, e, portanto, ser do tipo não

probabilística, a inferência estatística deve ser vista com cautela para a expansão dos resultados à população.

Descrição da análise dos dados

Os dados foram analisados por meio da Análise Fatorial Confirmatória (CFA) utilizando o pacote lavaan (Rosseel, 2012) do *software* R, para a validação da escala, e por meio do teste não paramétrico de Kruskal-Wallis, similar à análise de variância (Anova) quando os pressupostos paramétricos não são atendidos, para analisar a relação entre as variáveis “Desenvolvimento de competências” e “Experiência profissional dos avaliadores externos”. O tamanho do efeito (*effect size*) para esse tipo de teste é calculado pela estatística η^2 (eta ao quadrado). Testes sequenciais ao Kruskal-Wallis (*post-hoc*) foram executados com o objetivo de explorar as diferenças entre os níveis de experiência. Optou-se pelo teste de Dunn, apropriado para grupos com tamanhos diferentes, com ajuste de p valores pelo método BH, o qual controla a taxa de falsa detecção (Daniel, 2000).

A CFA é frequentemente utilizada para fornecer evidências da validade da medição e, assim, verificar se os dados empíricos estão ajustados ao modelo teórico da pesquisa. Esse método estatístico possibilita aos pesquisadores a análise das correlações entre as variáveis da pesquisa, com o intuito de determinar o grau em que essas correlações são consistentes com as expectativas teóricas (Badalos, 2018). Como os itens de cada competência foram respondidos em escala tipo Likert, portanto ordinal, utilizou-se o método de estimação denominado mínimos quadrados ponderados ajustados para média e variância (Distefano, 2002; Flora & Curran, 2004). Foi ainda utilizado, para verificação da confiabilidade da escala, o coeficiente ômega, considerado de maior acurácia que o coeficiente alpha (Badalos, 2018; Deng & Chan, 2017).

Na CFA, os índices de ajuste (*fit index*) utilizados para analisar a validade dos construtos desta pesquisa, como o Comparative Fit Index (CFI), o Standardized Root Mean Square Residual (SRMR), o Root Mean Square Error of Approximation (RMSEA) e o Qui-quadrado Normalizado, têm os padrões de referência abordados por Brown (2006), Kline (2016) e Badalos (2018) em seus estudos.

Com índices de ajuste adequados, os itens de cada uma das dimensões foram agrupados em uma única medida, a soma. Ainda que os participantes tenham respondido em uma escala ordinal, o uso da soma ou da média para agrupar itens de um fator (dimensão) de cada respondente é considerada a forma mais tradicional de medida de escore de um construto, apesar de existirem outras possibilidades. A soma alcança uma amplitude maior do que a escala original, tornando possível o uso da medida do construto em outras operações, como uma medida intervalar ou mesmo ordinal (Badalos, 2018).

ANÁLISE DOS DADOS

Os dados foram analisados considerando as duas fases da pesquisa, porém com apresentação em função do produto gerado ou do procedimento analítico. Assim, primeiramente apresenta-se a lista de competências obtidas após a Delphi, seguida do processo de validação da escala e, por último, da análise do desenvolvimento das competências e a relação com a experiência.

Lista de competências consideradas adequadas ao Sinaes

As competências em avaliação elencadas na Tabela 1 foram aquelas consideradas mais adequadas ao contexto do processo avaliativo do Sinaes na Delphi, realizada com os cinco especialistas, cuja maioria é homem, entre 41 e 50 anos e fez mais de quinze avaliações, todos com doutorado. Essas competências foram levadas a campo com os avaliadores externos para que não só o grau de concordância em relação a seu pleno desenvolvimento pudesse ser identificado, mas também para evidenciar a maneira como esse nível de concordância varia com a experiência profissional deles.

TABELA 1
Competências em avaliação adequadas ao Sinaes

1.0 - PRÁTICA PROFISSIONAL	
1.1	Atuar de forma ética e buscar integridade e honestidade na realização de avaliações.
1.2	Conhecer as teorias, modelos, métodos e ferramentas de avaliação e permanecer informado sobre o novo pensamento e melhores práticas de avaliação.
1.3	Aplicar padrões de avaliação profissional.
1.4	Respeitar clientes, entrevistados, participantes do programa e outras partes interessadas.
1.5	Identificar como a prática de avaliação pode promover a justiça social e o bem público.
1.6	Utilizar evidências de forma sistemática para fazer julgamentos avaliativos.
1.7	Demonstrar flexibilidade.
1.8	Defender o campo da avaliação e seu valor.
1.9	Estar comprometido com a transparência em todos os aspectos da avaliação.
1.10	Demonstrar credibilidade profissional, discricção e confidencialidade ao longo dos processos de avaliação.
1.11	Praticar meta-avaliação, buscando formas de incorporar a responsabilidade em seu trabalho.
1.12	Oferecer <i>feedback</i> formativo para sua equipe e colegas.
1.13	Ter sensibilidade, respeitar e cumprir os protocolos e práticas culturais.
1.14	Mostrar sensibilidade ética em contextos sociopolíticos específicos.
2.0 - METODOLOGIA	
2.1	Coletar, analisar, interpretar e sintetizar os dados da avaliação usando métodos e critérios apropriados.
2.2	Desenvolver um projeto de avaliação coerente com os objetivos da avaliação.
2.3	Compreender a base de conhecimento da avaliação (termos, conceitos, teorias, suposições) e revisar a literatura.

(continua)

(continuação)

2.0 - METODOLOGIA	
2.4	Identificar os pressupostos e usar a lógica, metodologia e teorias do programa de avaliação apropriadamente.
2.5	Discriminar tópicos e questões da avaliação.
2.6	Identificar fontes relevantes de evidências e procedimentos de amostragem, incluindo o escopo e a validade de cada fonte.
2.7	Avaliar a confiabilidade e a validade dos dados por meio do uso de verificações de dados, ensaios de controle e comparação, triangulação de resultados e análises cruzadas.
2.8	Identificar propósitos e necessidades de avaliação, justificando as decisões ao longo da avaliação.
2.9	Relatar, de vários ângulos, se as constatações da avaliação são credíveis, úteis e acionáveis, se respondem às perguntas feitas e se têm clareza e transparência quanto a metodologia, escolhas e interpretações avaliativas feitas.
2.10	Integrar sistematicamente evidências de várias fontes, identificando ligações causais, gerando descobertas válidas e observando pontos fortes, lacunas de dados e limitações nas avaliações.
2.11	Realizar meta-avaliações.
2.12	Fornecer pensamento crítico e habilidades analíticas e de síntese para a avaliação.
2.13	Compreender abordagens participativas e os contextos nos quais são mais apropriadas.
2.14	Perceber quais métodos são apropriados para fins formativos/somativos.
3.0 - ANÁLISE SITUACIONAL	
3.1	Identificar as necessidades das partes interessadas e sua capacidade de participar, reconhecendo, respeitando, respondendo a aspectos da diversidade e mantendo relacionamentos confiáveis.
3.2	Reunir a análise contextual e o engajamento, para que a evidência, a análise, a síntese e a interpretação da atividade avaliativa sejam credíveis e válidas para o leque de pessoas envolvidas e afetadas pela avaliação.
3.3	Comunicar processos de avaliação e resultados de maneira oportuna, apropriada e eficaz aos usuários pretendidos.
3.4	Identificar, entender, articular e levar em conta o contexto e a situação mais amplos e relevantes para a avaliação.
3.5	Promover o uso da avaliação e influência no contexto da avaliação.
3.6	Respeitar a singularidade do local de avaliação e do cliente.
3.7	Utilizar o entendimento contextual em análise, síntese, interpretação avaliativa e elaboração de relatórios.
3.8	Identificar e incorporar protocolos culturais apropriados para interagir com a comunidade, incluindo a incorporação de conhecimentos culturais na equipe de avaliação.
3.9	Identificar barreiras e facilitadores que influenciarão o esforço de avaliação.
3.10	Articular como as relações de poder podem influenciar processos ou resultados específicos de avaliação.
4.0 - GERENCIAMENTO DE PROJETO DE AVALIAÇÃO	
4.1	Identificar os recursos necessários para a avaliação, como informações, conhecimento pessoal, instrumentos e financeiros.
4.2	Selecionar e utilizar efetivamente ferramentas de avaliação apropriadas e adequadas à finalidade.
4.3	Definir parâmetros de trabalho, planos e acordos formais para a avaliação.
4.4	Utilizar habilidades de gerenciamento e facilitação de equipes de avaliação.
4.5	Estabelecer comunicação com os clientes durante todo o processo de avaliação.
4.6	Documentar processos e produtos de avaliação.
4.7	Conduzir a avaliação de maneira ininterrupta.
4.8	Terminar a avaliação em tempo hábil.

(continua)

(continuação)

4.0 - GERENCIAMENTO DE PROJETO DE AVALIAÇÃO	
4.9	Responder às propostas de avaliação.
4.10	Gerenciar e proteger os dados de avaliação.
4.11	Identificar e mitigar problemas / questões de maneiras construtivas e úteis.
4.12	Relatar ao público de forma justa e no interesse público o que foi acordado.
5.0 - PRÁTICA REFLEXIVA	
5.1	Ter consciência de si mesmo como avaliador (conhecimentos, habilidades, atitudes).
5.2	Refletir sobre a prática de avaliação (competências e áreas pessoais de crescimento).
5.3	Construir relacionamentos profissionais para aprimorar a prática de avaliação.
5.4	Reconhecer e ser transparentes sobre sua influência no processo de avaliação.
6.0 - COMPETÊNCIA INTERPESSOAL	
6.1	Facilitar a interação interpessoal construtiva (trabalho em equipe, facilitação de grupo, processamento).
6.2	Demonstrar habilidades na comunicação verbal e não verbal eficaz, apropriada e respeitosa.
6.3	Utilizar habilidades de resolução de conflitos.
6.4	Usar habilidades e tecnologias de comunicação escrita na prática de avaliação.
6.5	Demonstrar competência transcultural.
6.6	Criar confiança ao longo da avaliação.
6.7	Escutar, entender e engajar diferentes perspectivas.
6.8	Manter uma perspectiva objetiva sobre avaliação.
6.9	Demonstrar consciência de gênero e sensibilidade cultural.

Fonte: Elaboração dos autores a partir do resultado da Delphi.

A análise genérica da Tabela 1 possibilita dizer que as 63 competências resultantes da *Delphi* e sua distribuição entre as seis dimensões teóricas demonstram sua relativa conformidade numérica e conceitual com as primeiras pesquisas sobre competências em avaliação feitas por King et al. (2001). Contudo, ressalta-se que essa análise aconteceu de forma pouco detalhada em relação às competências de cada dimensão teórica ou grupo de competências específicas, uma vez que não foram encontrados na literatura brasileira e nem em plataformas virtuais do Inep estudos que abordassem competências essenciais para os avaliadores do Sinaes, o que impossibilita a existência de parâmetros para a comparação de resultados, bem como reduz a margem de diálogo dos achados da Delphi com a literatura sobre competências em avaliação.

Validação da escala

Os dados coletados com os avaliadores a respeito do desenvolvimento das competências resultantes da Delphi foram analisados por dimensão. Os principais índices de ajuste (*fit-index*) do modelo estão elencados na Tabela 2.

TABELA 2
Indicadores de ajuste do modelo (*fit-index*)

	OBTIDO	PARÂMETROS DE ACEITAÇÃO A PARTIR DA LITERATURA
Qui-Quadrado (<i>Chi-Square</i>)	3027,223	---
Df (<i>Degrees of freedom</i>)	1875,000	---
p-value	< 0,001	> 0,05
Qui-Quadrado/Df (Qui-Quadrado Normalizado)	1,615	< 2
CFI	0,974	≥ 0,95
SRMR	0,056	≤ 0,05
RMSEA	0,034	≤ 0,05
RMSEA - limite inferior	0,032	---
RMSEA - limite superior	0,036	---

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

Os resultados evidenciam que as variáveis observadas, traduzidas nas 63 competências em avaliação levadas a campo, conseguem medir os respectivos construtos, representados por cada dimensão teórica às quais estão relacionadas.

Os resultados apresentados na Tabela 3 referem-se ao valor obtido do coeficiente ômega a respeito da confiabilidade e à amplitude das cargas fatoriais padronizadas de cada dimensão. O grau em que os itens compartilham a variância com os construtos é medido por suas cargas fatoriais. Os valores dessas cargas são diferentes entre os itens porque alguns itens medem melhor uma variável latente do que outros, ou seja, variáveis com cargas fatoriais de valores relativamente altos indicam que essas variáveis se relacionam fortemente com seus respectivos construtos (Badalos, 2018). Por outro lado, o coeficiente ômega verifica se o número de itens está adequado para medir os construtos da pesquisa (Deng & Chan, 2017). Conclui-se que os valores obtidos estão dentro dos padrões de aceitabilidade para a validação da escala. Os valores individuais das cargas fatoriais, bem como das correlações (que variam entre 0,70 e 0,90), podem ser visualizados no Apêndice deste artigo.

TABELA 3
Dados de confiabilidade e cargas fatoriais

DIMENSÃO (NÚMERO DE ITENS)	COEFICIENTE ÔMEGA	AMPLITUDE DA CARGA FATORIAL (PADRONIZADA)
Prática Profissional (14)	0,84	0,625 - 0,903
Metodologia (14)	0,89	0,677 - 0,873
Análise Situacional (10)	0,91	0,687 - 0,886

(continua)

(continuação)

DIMENSÃO (NÚMERO DE ITENS)	COEFICIENTE ÔMEGA	AMPLITUDE DA CARGA FATORIAL (PADRONIZADA)
Gerenciamento de Projeto de Avaliação (12)	0,89	0,720 – 0,875
Prática Reflexiva (4)	0,82	0,802 – 0,938
Competência Interpessoal (9)	0,90	0,776 – 0,903

Fonte: Elaboração dos autores.

Análise do desenvolvimento das competências e sua relação com a experiência

A média e a mediana das respostas por docente em cada uma das dimensões teóricas aproximam-se do nível 4, valor máximo da escala (ver Tabela 4). Isso indica que a grande maioria dos respondentes da amostra concorda que tem as competências elencadas em todas as dimensões teóricas plenamente desenvolvidas.

TABELA 4

Medidas descritivas para a resposta média por docente segundo dimensão teórica*

DIMENSÃO	Nº DE ITENS	MÉDIA*	DESVIO-PADRÃO	MÍNIMO	MEDIANA	MÁXIMO
Prática Profissional	14	3,73	0,28	2,57	3,79	4,00
Metodologia	14	3,50	0,43	1,93	3,57	4,00
Análise Situacional	10	3,58	0,43	1,90	3,70	4,00
Gerenciamento de Projeto de Avaliação	12	3,64	0,40	1,75	3,75	4,00
Prática Reflexiva	4	3,75	0,38	2,00	4,00	4,00
Competência Interpessoal	9	3,68	0,38	2,00	3,89	4,00

Fonte: Elaboração dos autores a partir de pesquisa de campo (2021).

* n = 529.

As dimensões Metodologia e Análise Situacional são as que têm os maiores desvios-padrão, o que indica que as respostas estão mais distribuídas entre as variações da escala. Por outro lado, nas dimensões Prática Reflexiva e Competência Interpessoal, os desvios-padrão foram os menores, o que indica que as respostas estão mais concentradas no grau máximo da escala. Um maior ou menor desvio-padrão pode indicar aquelas dimensões em que o relato de desenvolvimento de competências varia mais entre os respondentes, ou seja, há mais avaliadores que se consideram plenamente competentes e também competentes. Como as médias ficaram próximas do valor máximo, não se pode afirmar que os avaliadores se julgaram com alguma competência não desenvolvida.

A Tabela 5 apresenta as estatísticas da média das respostas por dimensão, abertas em função da experiência em avaliação dos respondentes. Em todas as dimensões é possível perceber que a média aumenta à medida que a experiência aumenta, enquanto o desvio-padrão diminui. A mediana também aumenta, nota-

damente da categoria Nenhuma para as categorias De 4 a 9 e Mais de 10, exceto para a dimensão Prática Reflexiva.

Para verificar se as diferenças entre grupos de experiência são significativas, foram conduzidos testes não paramétricos (Kruskal-Wallis) para cada dimensão. Vale ressaltar que os procedimentos tradicionais de testes paramétricos, como o da Anova, pressupõem, entre outros aspectos, normalidade na distribuição dos dados e homogeneidade de variâncias (Hair et al., 2009). No caso em questão, para todas as dimensões, nenhum dos dois aspectos foi atendido. Por isso optou-se por proceder com o teste não paramétrico de Kruskal-Wallis (Daniel, 2000)

TABELA 5
Experiência e competências (média das respostas por dimensão)

DIMENSÃO	EXPERIÊNCIA (Nº DE AVALIAÇÕES)	N	MÉDIA	DESVIO-PADRÃO	MEDIANA	MÍN.	MÁX.	X ²	η ²
Prática Profissional	Nenhuma	152	3,66	0,34	3,79	2,57	4	13,2*	1,9%
	De 1 a 3	153	3,71	0,27	3,79	2,86	4		
	De 4 a 9	171	3,79	0,23	3,86	2,64	4		
	Mais de 10	53	3,80	0,20	3,86	3,07	4		
Metodologia	Nenhuma	152	3,37	0,51	3,43	1,93	4	17,8*	2,8%
	De 1 a 3	153	3,48	0,41	3,57	2,64	4		
	De 4 a 9	171	3,58	0,36	3,71	2,50	4		
	Mais de 10	53	3,64	0,34	3,71	2,79	4		
Análise Situacional	Nenhuma	152	3,44	0,50	3,55	2,00	4	22,2*	3,6%
	De 1 a 3	153	3,57	0,41	3,70	1,90	4		
	De 4 a 9	171	3,66	0,36	3,80	2,40	4		
	Mais de 10	53	3,71	0,36	3,90	2,30	4		
Gerenciamento de Projeto de Avaliação	Nenhuma	152	3,50	0,46	3,67	2,00	4	27,3*	4,6%
	De 1 a 3	153	3,62	0,40	3,75	1,75	4		
	De 4 a 9	171	3,73	0,33	3,83	2,67	4		
	Mais de 10	53	3,77	0,31	3,92	2,92	4		
Prática Reflexiva	Nenhuma	152	3,66	0,44	4,00	2,00	4	14,7*	2,2%
	De 1 a 3	153	3,75	0,36	4,00	2,50	4		
	De 4 a 9	171	3,8	0,34	4,00	2,5	4		
	Mais de 10	53	3,86	0,28	4,00	3,00	4		
Competência Interpessoal	Nenhuma	152	3,58	0,45	3,78	2,00	4	104*	1,4%
	De 1 a 3	153	3,67	0,35	3,78	2,78	4		
	De 4 a 9	171	3,73	0,34	3,89	2,67	4		
	Mais de 10	53	3,79	0,26	3,89	3,00	4		

Fonte: Elaboração dos autores.

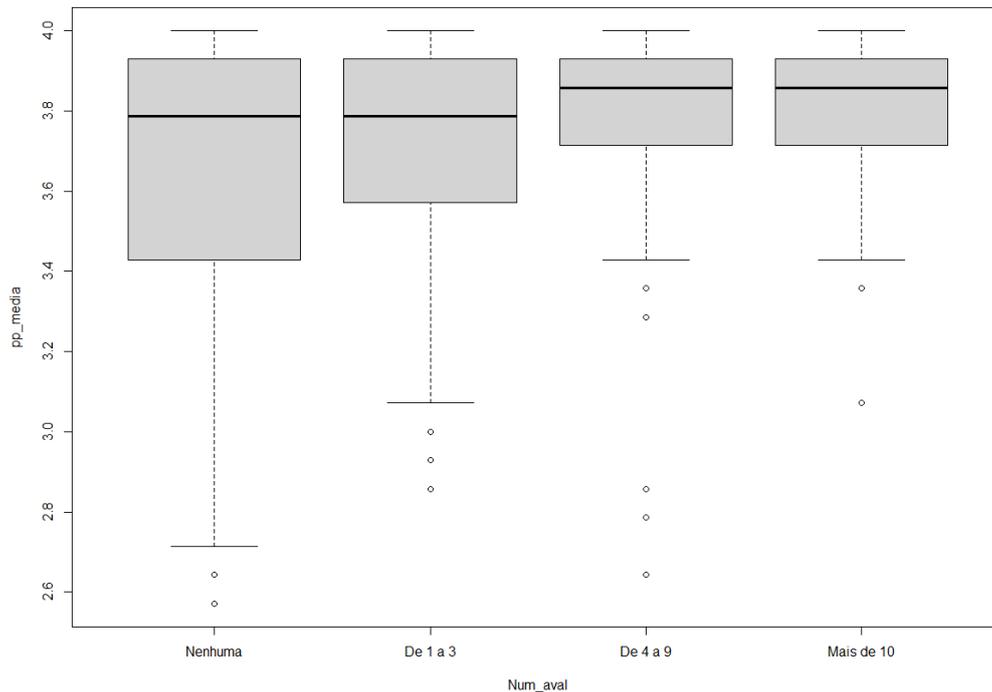
* p-valor < 0,001.

Mesmo sendo possível verificar que, à medida que a experiência em avaliação aumenta, as médias das competências por dimensão também aumentam, o tamanho do efeito (*effect size*), calculado a partir do η^2 , é considerado relativamente baixo. A experiência em avaliação explica entre 1,4% (Competência Interpessoal) e 4,6% (Gerenciamento de Projeto de Avaliação) da variância das competências. Ou seja, existem outras variáveis que explicam o percentual restante do relato de desenvolvimento das competências dos avaliadores respondentes e que não foram objeto desta pesquisa. Isso pode ter acontecido pois os valores dos métodos estatísticos de ajuste global que parecem sugerir ajuste adequado não indicam que o poder explicativo do modelo é alto para resultados individuais (Kline, 2016). Ainda assim, entende-se que tais valores do tamanho do efeito não devem ser negligenciados ou entendidos como irrelevantes.

Os dados da Prática Profissional podem ser visualizados na Figura 1 (*boxplot*²), que ilustra a dispersão das respostas nas escalas de concordância sobre o desenvolvimento de competências com os grupos relacionados à quantidade de avaliações realizadas. O gráfico permite uma interpretação visual do efeito da experiência em avaliação em relação ao desenvolvimento de competências nessa dimensão, evidenciando a forma como grupos com mais experiência, notadamente mais de 4 avaliações (categoria de 4 a 9 avaliações), distinguem-se do grupo com nenhuma avaliação, ainda que todos os grupos tenham obtido médias de desenvolvimento de competências próximas do ponto mais elevado de cada escala. Os gráficos das demais dimensões apresentam quase todos o mesmo comportamento (ver Apêndice). O gráfico da dimensão Prática Reflexiva é o único levemente diferente, pois a mediana atinge o ponto máximo, mas, por outro lado, a redução visualizada da variância é observada como nos demais.

2 O *boxplot*, ou diagrama de caixa, é um gráfico que representa a variação observada nos dados de uma variável numérica por meio de quartis. A caixa central representa onde estão 50% das observações, localizadas entre os 1º e 3º quartis. O traço que divide a caixa indica a mediana ou 2º quartil. Ao longo das caudas estão dispersos o restante dos dados, exceto os valores mais discrepantes (*outliers*), representados por pontos abaixo da cauda inferior (ou acima da superior).

FIGURA 1
Boxplot da Experiência e Prática Profissional



Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

Após o teste de Kruskal-Wallis, considerando que os resultados indicaram haver diferença significativa entre os grupos, foram conduzidos testes *post hoc* (teste de Dunn). Testes *post hoc* examinam as variáveis dependentes entre todos os possíveis pares de grupos (Hair et al., 2009). O teste de Dunn tem o intuito de comparar grupos aos pares e verificar se as distribuições são similares ou não (Daniel, 2000). Como o teste é feito a partir das médias dos postos (valores ordenados), os resultados apresentados na Tabela 6 referem-se à diferença entre os valores observados na comparação aos pares.

TABELA 6
Resultados no teste de Dunn entre pares de níveis de experiência (nível b - nível a)

PAR DE NÍVEIS DE EXPERIÊNCIA			DIMENSÃO					
Nível a	vs	Nível b	PP	MET	AS	GPA	PR	CI
Nenhuma		De 1 a 3	3,64	22,04	29,74	34,80	25,57	20,26
		De 4 a 9	49,37*	59,53**	67,24***	75,95***	48,80**	45,49*
		Mais de 10	52,99*	78,39**	87,90***	95,35***	64,40**	59,11*
De 1 a 3		De 4 a 9	45,73*	37,49*	37,50*	41,14*	23,23	25,23
		Mais de 10	49,35*	56,35*	58,16*	60,55*	38,84	38,85
De 4 a 9		Mais de 10	3,62	18,86	20,66	19,40	15,60	13,62

Fonte: Elaboração dos autores (2021).

* p-valor < 0,05; ** p-valor < 0,01; *** p-valor < 0,001 nos casos em a diferença das estimativas dif = b - a > 0.

Pode-se verificar, a partir dos dados da Tabela 6, que, em todas as dimensões, a diferença entre não ter experiência com o Sinaes e ter alguma experiência aumenta, notadamente, a partir de mais de 4 avaliações, indicando que o relato acerca do desenvolvimento das competências em avaliação é maior. Ou seja, avaliadores com mais de 4 avaliações (categoria De 4 a 9 avaliações) relatam ter competências desenvolvidas num grau maior do que aqueles que não realizaram nenhuma avaliação.

Desse modo, a aplicação dos testes *post hoc* evidencia que a experiência de alguma forma interfere no relato do desenvolvimento de competência do avaliador. Além disso, evidencia também que, à medida que a experiência do avaliador – representada pelo número de avaliações feitas por ele – aumenta, maior é seu grau de concordância em relação ao pleno desenvolvimento dessas competências.

Os resultados vão ao encontro de parte substancial da literatura de avaliação, que enfatiza a importância da experiência prática para os avaliadores, uma vez que estes não conseguem desenvolver as competências necessárias para executar as avaliações se não tiverem tais experiências (Kaesbauer, 2012). Nesse sentido, experiências práticas também contribuem para o desenvolvimento de habilidades necessárias para o sucesso profissional de um avaliador (Trevisan, 2002).

CONSIDERAÇÕES FINAIS

Os avaliadores do Sinaes devem estar cientes de que o pleno desenvolvimento de competências em avaliação é relevante para a execução eficaz das avaliações nas instituições de ensino superior do país e de que sua experiência na execução de avaliações pode contribuir para que esse desenvolvimento aconteça de maneira eficiente, gerando a oferta de serviços de qualidade pelas IES e proporcionando benefícios à coletividade. Por isso, esta pesquisa teve como objetivo geral analisar a relação existente entre o desenvolvimento de competências em avaliação autodeclarado pelos avaliadores externos do Sinaes e sua experiência na execução de avaliações.

No que diz respeito ao desenvolvimento de competências, os dados empíricos mostraram que, em sua grande maioria, os avaliadores externos concordaram totalmente que apresentam desenvolvidas ou plenamente desenvolvidas as competências em avaliação listadas e entendidas como contextualizadas ao Sinaes.

A análise da relação entre a experiência em avaliação e o desenvolvimento de competências evidenciou que, em todas as dimensões teóricas de competências em avaliação, quanto maior a quantidade de avaliações feitas, maior foi a confiança do avaliador em relatar que tem as competências desenvolvidas. Esse resultado aponta, empiricamente, que há uma relação entre as duas variáveis. Contudo os testes estatísticos também mostraram que o tamanho do efeito da experiência sobre o relato de desenvolvimento de competências é relativamente baixo, o que significa que existem outras variáveis que explicam o desenvolvimento das competências em avaliação

e que não fizeram parte do escopo desta pesquisa. Porém, apesar de baixo, ele não deve ser negligenciado. Uma implicação prática dos resultados encontrados refere-se à distribuição das equipes de avaliadores que trabalharão nas instituições. Como recomendação, o Inep deve, sempre que possível, integrar os avaliadores recém-incorporados ao BASis a avaliadores mais experientes, de preferência aqueles com mais de dez avaliações já realizadas ou, pelo menos, com mais de quatro avaliações concluídas. Essa estratégia pode contribuir para o desenvolvimento das competências dos novos avaliadores e, em última instância, produzir melhores avaliações.

O estudo sobre competências essenciais do avaliador proposto por King et al. (2001) representou uma tentativa de mostrar que é possível criar uma lista de competências específicas aplicáveis ao contexto da avaliação. Essa lista também pode constituir uma ferramenta útil de desenvolvimento profissional nos programas de avaliação, servindo como um elemento propulsor do envolvimento dos avaliadores em reflexões e discussões coletivas focadas em um rol de competências mais associadas à prática da avaliação (Stevahn et al., 2005b).

As listas de competências elaboradas por AEA (2018), CES (2018), AES (2013), EES (2011), UKES (2012) e Anzea (2011) representaram buscas pelo aprimoramento e ampliação dos estudos já realizados sobre as competências do avaliador. Diversas abordagens legitimam listas de competências em avaliação como essas, fazendo com que muitas associações atuem como plataformas para um diálogo interdisciplinar sobre “o que é necessário” para ser um avaliador competente, uma vez que elas são solicitadas a promover e melhorar a teoria, prática, entendimento e utilização de avaliação e sua contribuição para a sociedade (EES, 2011).

Como contribuição teórica, esta pesquisa pode estimular o surgimento ou crescimento de estudos que abordem a relação entre experiência em avaliação e desenvolvimento de competências no âmbito de cada dimensão teórica de competência em avaliação, notadamente com abordagens experimentais ou quase experimentais.

Vale ressaltar que este é um estudo do tipo observacional ou correlacional (Badalos, 2018) e, ainda que exista uma correlação entre as variáveis “experiência em avaliação” e “relato de desenvolvimento de competência”, isso não determina uma relação de causalidade entre elas. A existência de covariância sistemática entre causa e efeito é necessária, porém não suficiente, para estabelecer tal causalidade (Hair et al., 2009).

REFERÊNCIAS

- American Evaluation Society (AEA). (2018). *The 2018 AEA-evaluator competencies*. AEA.
<https://www.eval.org/Portals/0/Docs/AEA%20Evaluator%20Competencies.pdf>
- Aotearoa New Zealand Evaluation Association (Anzea). (2011). *Evaluator competencies*. Anzea.
https://anzea.org.nz/assets/Key-ANZEA-Files/110801_anzea_evaluator_competencies_final.pdf

- Australasian Evaluation Society (AES). (2013). *Evaluator's Professional Learning Competency Framework*. AES. https://www.aes.asn.au/images/AES_Evaluators_Competency_Framework.pdf
- Badalos, D. L. (2018). *Measurement theory and applications for the social sciences*. Guilford.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. Guilford.
- Canadian Evaluation Society (CES). (2018). *Competencies for Canadian evaluation practice*. CES. https://evaluationcanada.ca/files/pdf/2_competencies_cdn_evaluation_practice_2018.pdf
- Castro, L. C. da S. (2021). *Competências e experiência em avaliação: Uma análise dos avaliadores externos do Sinaes* [Dissertação de mestrado profissional, Universidade Federal da Bahia]. Repositório Institucional da UFBA. <https://repositorio.ufba.br/handle/ri/35525>
- Dalkey, N., & Helmer, O. (1962). *An experimental application of the Delphi method to the use of experts*. The Rand Corporation. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2009/RM727.1.pdf
- Daniel, W. W. (2000). *Applied nonparametric statistics* (2nd ed.). Cengage Learning.
- Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients Alpha and Omega. *Educational and Psychological Measurement*, 77(2), 185-203. <https://doi.org/10.1177/0013164416658325>
- Diaz, J., Chaudhary, A. K., Jayaratne, K. S. U., & Assan, E. (2020). Expanding evaluator competency research: Exploring competencies for program evaluation using the context of non-formal education. *Evaluation and Program Planning*, 79, Article e101790. <https://doi.org/10.1016/j.evalprogplan.2020.101790>
- Diretoria de Avaliação da Educação Superior (Daes). (2021). *Pedido de informação ao Inep registrado sob o protocolo n. 23546-060999/2021*. <http://www.consultaesic.cgu.gov.br/busca/SitePages/Principal.aspx>
- Distefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, 9(3), 327-346. https://doi.org/10.1207/S15328007SEM0903_2
- European Evaluation Society (EES). (2011). *EES Evaluation Capabilities Framework*. EES.
- Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466-491.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Análise multivariada de dados* (6a ed.). Bookman.
- Halaydna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.
- Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation*, 12, Article 10. <https://doi.org/10.7275/pdz9-th90>
- Kaesbauer, S. A. M. (2012). *Teaching evaluator competencies: An examination of doctoral programs* [PhD dissertation, University of Tennessee]. Tennessee Research and Creative Exchange. https://trace.tennessee.edu/utk_graddiss/1314/
- King, J. A., & Stevahn, L. (2015). Competencies for program evaluators in light of adaptive action: What? So what? Now what? Competencies for Program Evaluators in Light of Adaptive Action. *New Directions for Evaluation*, (145), 21-37. <https://doi.org/10.1002/ev.20109>
- King, J. A., Stevahn, L., Ghore, G., & Minnema, J. (2001). Toward a taxonomy of essential evaluator competencies. *American Journal of Evaluation*, 22(2), 229-247. <https://doi.org/10.1177/109821400102200206>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford.
- Portaria Normativa n. 840, de 24 de agosto de 2018. (2018). Dispõe sobre os procedimentos de competência do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira –

Inep referentes à avaliação de instituições de educação superior, de cursos de graduação e de desempenho acadêmico de estudantes. https://download.inep.gov.br/educacao_superior/avaliacao_institucional/legislacao_normas/2018/portaria_normativa_GM-MEC_n840_de_24082018.pdf

- Rosseel, Y. (2012). Ivaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>
- Stevahn, L., King, J., Ghere, G., & Minnema, J. (2005a). Evaluator competencies in university-based evaluation training programs. *The Canadian Journal of Program Evaluation*, 20(2), 101-123. <https://doi.org/10.3138/cjpe.20.006>
- Stevahn, L., King, J., Ghere, G., & Minnema, J. (2005b). Establishing essential competencies for program evaluators. *American Journal of Evaluation*, 26(1), 43-59. <https://doi.org/10.1177/1098214004273180>
- Trevisan, M. S. (2002). Enhancing practical evaluation training through long-term evaluation projects. *American Journal of Evaluation*, 23(1), 81-92. [https://doi.org/10.1016/S1098-2140\(01\)00163-1](https://doi.org/10.1016/S1098-2140(01)00163-1)
- Trevisan, M. S. (2004). Practical training in evaluation: A review of the literature. *American Journal of Evaluation*, 25(2), 255-272. <https://doi.org/10.1016/j.ameval.2004.03.002>
- UK Evaluation Society (UKES). (2012). *UK Evaluation Society Framework of Evaluation Capabilities*. UKES. <https://www.evaluation.org.uk/app/uploads/2019/04/UK-Evaluation-Society-Framework-of-Evaluation-Capabilities.pdf>
- Wehipeihana, N., Bailey, R., Davidson, E. J., & McKegg, K. (2014). Evaluator competencies: The Aotearoa New Zealand Experience. *The Canadian Journal of Program Evaluation*, 28(3), 49-69. <https://doi.org/10.3138/cjpe.0028.007>

APÊNDICE

TABELA A1

Dados estatísticos da dimensão Prática Profissional

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)				
				1	2	3	4	
Prática Profissional	1	0,84	0,713	1 (0,19)	1 (0,19)	28 (5,29)	499 (94,33)	
	2		0,711		12(2,27)	233(44,04)	284(53,69)	
	3		0,713		10(1,89)	180(34,03)	339(64,08)	
	4		0,790			37(6,99)	492(93,01)	
	5		0,626		27(5,10)	140(26,47)	362(68,43)	
	6		0,746	1(0,19)	13(2,46)	148(27,98)	367(69,38)	
	7		0,625	5(0,95)	31(5,86)	180(34,03)	313(59,17)	
	8		0,714	2(0,38)	9(1,70)	148(27,98)	370(69,94)	
	9		0,822		2(0,38)	41(7,75)	486(91,87)	
	10		0,903			30(5,67)	499(94,33)	
	11		0,683	2(0,38)	20(3,78)	156(29,49)	351(66,35)	
	12		0,782	2(0,38)	15(2,84)	165(31,19)	347(65,60)	
	13		0,791			3(0,57)	84(15,88)	442(83,55)
	14		0,852		4(0,76)	101(19,09)	424(80,15)	

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A2**Dados estatísticos da dimensão Metodologia**

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)			
				1	2	3	4
Metodologia	1	0,89	0,784		19(3,59)	166(31,38)	344(65,03)
	2		0,788		14(2,65)	188(35,54)	327(61,81)
	3		0,777		23(4,35)	210(39,70)	296(55,95)
	4		0,829		20(3,78)	207(39,13)	302(57,09)
	5		0,788	1(0,19)	19(3,59)	185(39,97)	324(61,25)
	6		0,760		24(4,54)	201(38)	304(57,47)
	7		0,694	5(0,95)	75(14,18)	252(47,64)	197(37,24)
	8		0,873	1(0,19)	13(2,46)	188(35,54)	327(61,81)
	9		0,832	2(0,38)	24(4,54)	231(43,67)	272(51,42)
	10		0,850		39(7,37)	211(39,89)	279(52,74)
	11		0,677	3(0,57)	69(13,04)	260(49,15)	197(37,24)
	12		0,864		15(2,84)	167(31,57)	347(65,60)
	13		0,851		21(3,97)	193(36,48)	315(59,55)
	14		0,817		39(7,37)	196(37,05)	294(55,58)

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A3**Dados estatísticos da dimensão Análise Situacional**

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)			
				1	2	3	4
Análise Situacional	1	0,91	0,818	1(0,19)	11(2,08)	137(25,90)	380(71,83)
	2		0,879	1(0,19)	13(2,46)	197(37,24)	318(60,11)
	3		0,806	5(0,95)	19(3,59)	152(28,73)	353(66,73)
	4		0,766	6(1,13)	30(5,67)	200(37,81)	293(55,39)
	5		0,870		12(2,27)	185(34,97)	332(62,76)
	6		0,815	1(0,19)	14(2,65)	119(22,50)	395(74,67)
	7		0,886		16(3,02)	168(31,76)	345(65,22)
	8		0,765	1(0,19)	40(7,56)	205(38,75)	283(53,50)
	9		0,814		26(4,91)	201(38)	302(57,09)
	10		0,687	5(0,95)	35(6,62)	174(32,89)	315(59,55)

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A4**Dados estatísticos da dimensão Gerenciamento de Projeto de Avaliação**

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)			
				1	2	3	4
Gerenciamento de Projeto de Avaliação	1	0,89	0,774	1(0,19)	23(4,35)	162(30,62)	343(64,84)
	2		0,828	3(0,57)	19(3,59)	172(32,51)	335(63,33)
	3		0,854	4(0,76)	10(1,89)	175(33,08)	340(64,27)
	4		0,841	2(0,38)	25(4,73)	178(33,65)	324(61,25)
	5		0,818	1(0,19)	19(3,59)	142(26,84)	367(69,38)
	6		0,859	1(0,19)	10(1,89)	147(27,79)	371(70,13)
	7		0,745		33(6,24)	162(30,62)	334(63,14)
	8		0,720	2(0,38)	15(2,84)	148(27,98)	364(68,81)
	9		0,875		15(2,84)	148(27,98)	366(69,19)
	10		0,822		8(1,51)	86(16,26)	435(82,23)
	11		0,861	1(0,19)	6(1,13)	172(32,51)	350(66,16)
	12		0,752	8(1,51)	16(3,02)	145(27,41)	360(68,05)

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A5**Dados estatísticos da dimensão Prática Reflexiva**

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)			
				1	2	3	4
Prática Reflexiva	1	0,82	0,902		3(0,57)	105(19,85)	421(79,58)
	2		0,938		3(0,57)	124(23,44)	402(75,99)
	3		0,822		12(2,27)	126(23,82)	391(73,91)
	4		0,802	3(0,57)	9(1,70)	102(19,28)	415(78,45)

Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A6**Dados estatísticos da dimensão Competência Interpessoal**

DIMENSÃO	NÚMERO ITEM	COEFICIENTE ÔMEGA	CARGA FATORIAL (PADRONIZADA)	FREQUÊNCIAS (%)			
				1	2	3	4
Competência Interpessoal	1	0,90	0,849		5(0,95)	140(26,47)	384(72,59)
	2		0,832		4(0,76)	120(22,68%)	405(76,56)
	3		0,826		13(2,46)	184(34,78)	332(62,76)
	4		0,818		7(1,32)	140(26,47)	382(72,21)
	5		0,778	5(0,95)	30(5,67)	213(40,26)	281(53,12)
	6		0,900		3(0,57)	139(26,28)	387(73,16)
	7		0,903		11(2,08)	134(25,33)	384(72,59)
	8		0,867		2(0,38)	136(25,71)	391(73,91)
	9		0,776	2(0,38)	14(2,65)	130(24,57)	383(72,40)

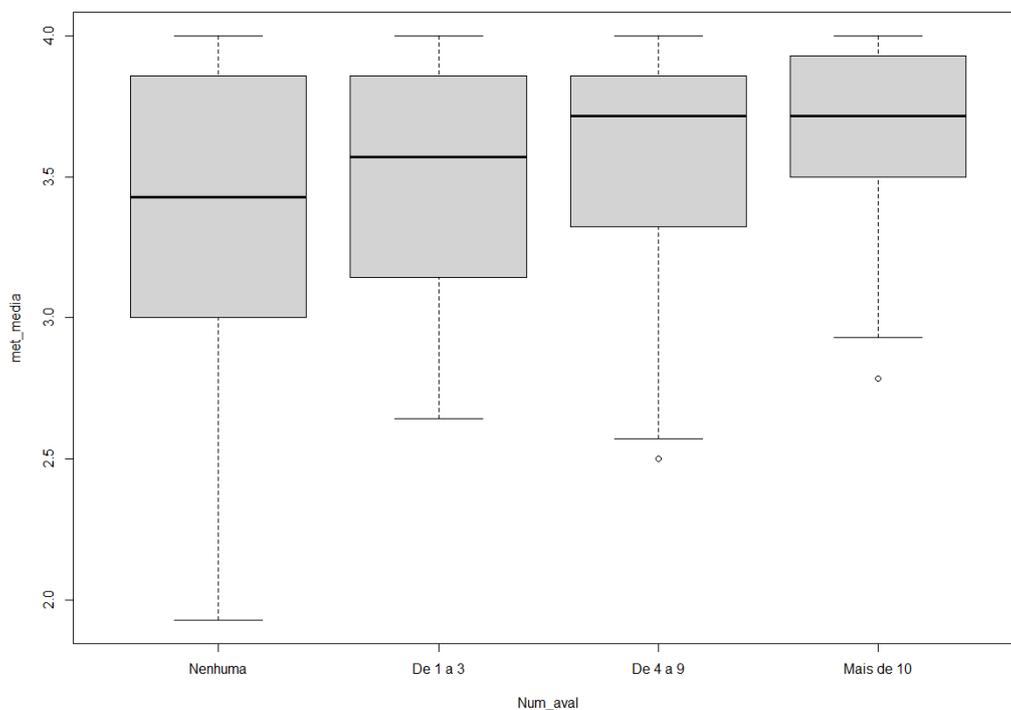
Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A7
Matriz de correlação entre as dimensões teóricas

	PP	MET	AS	GPA	PR	CI
PP	1					
MET	0,843	1				
AS	0,781	0,842	1			
GPA	0,737	0,831	0,893	1		
PR	0,736	0,708	0,833	0,835	1	
CI	0,741	0,750	0,853	0,869	0,903	1

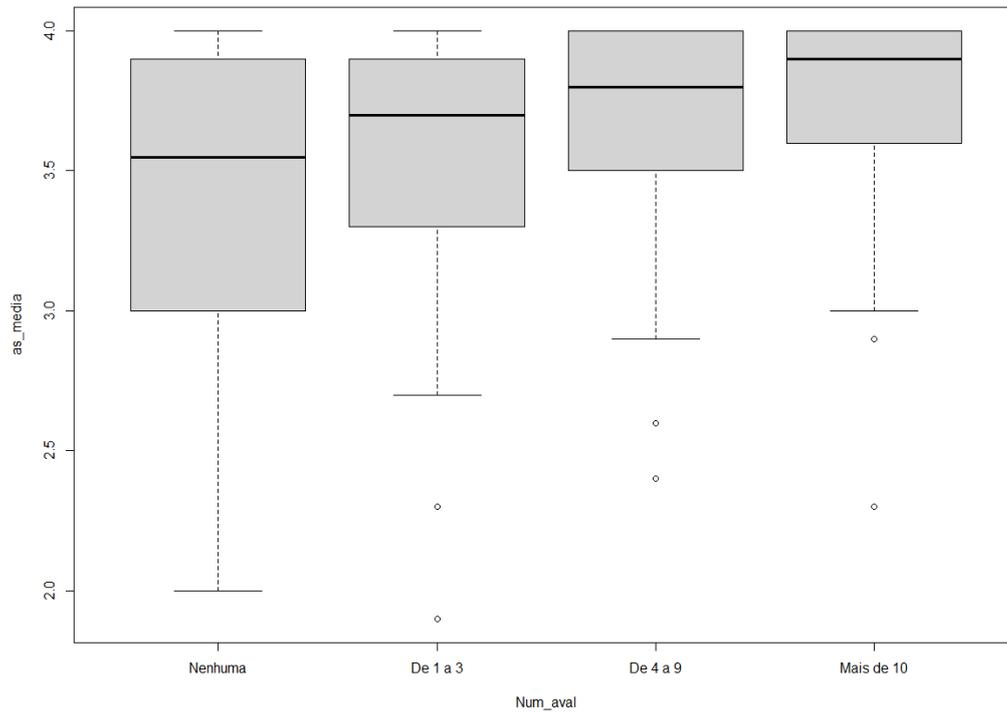
Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

FIGURA A1
Experiência e Metodologia



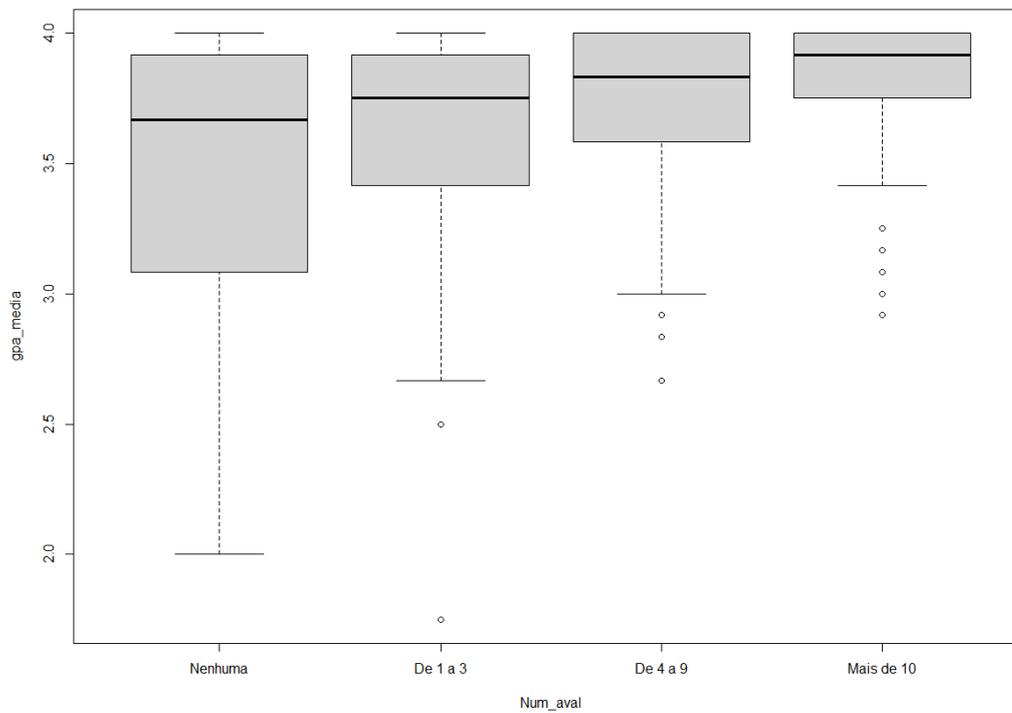
Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

FIGURA A2
Experiência e Análise Situacional



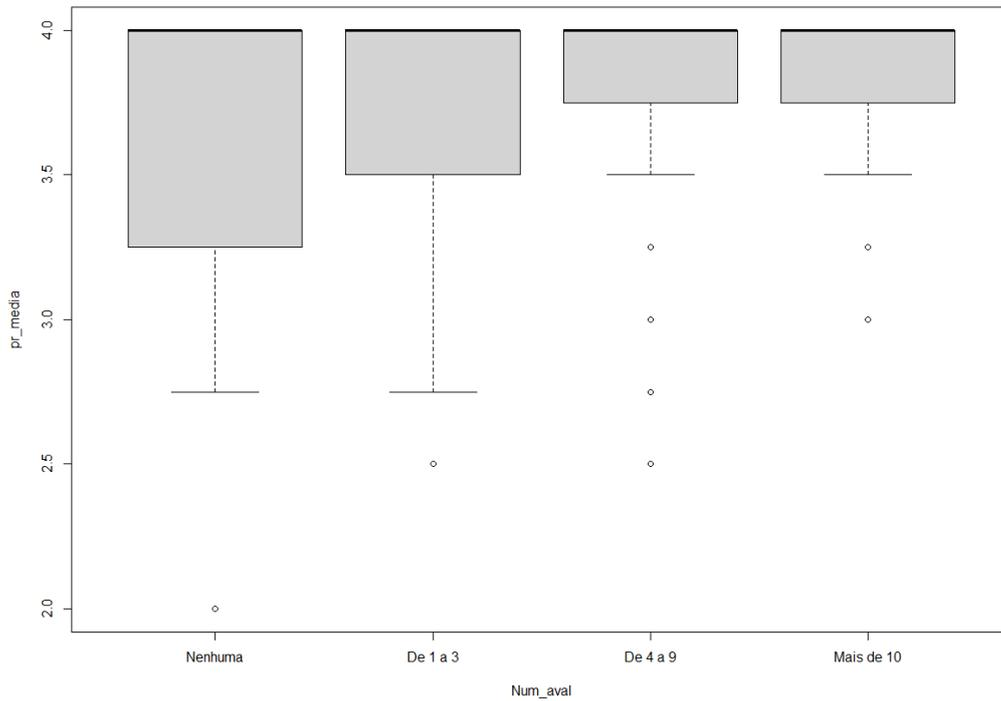
Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

FIGURA A3
Experiência e Gerenciamento de Projeto de Avaliação



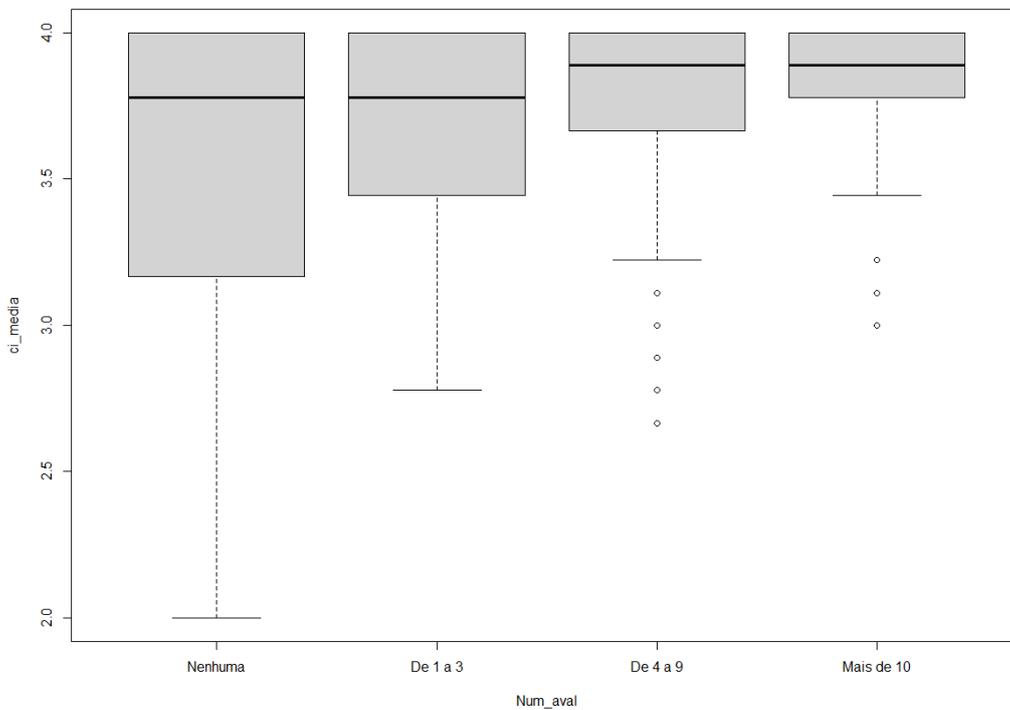
Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

FIGURA A4
Experiência e Prática Reflexiva



Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

FIGURA A5
Experiência e Competência Interpessoal



Fonte: Elaboração dos autores a partir dos resultados da pesquisa de campo.

TABELA A8
Matriz comparativa das listas de competências em avaliação

DIMENSÃO	COMPETÊNCIAS	KING	AEA	CES	AES	ANZEA	EES	UKES
Prática Profissional	1.1. Atuar de forma ética e buscar integridade e honestidade na realização de avaliações.	X	X	X	X	X	X	X
	1.2. Conhecer as teorias, modelos, métodos e ferramentas de avaliação, e permanecer informado sobre o novo pensamento e melhores práticas de avaliação.	X	X	X	X		X	X
	1.3. Aplicar padrões de avaliação profissional.	X	X	X	X	X		
	1.4. Respeitar clientes, entrevistados, participantes do programa e outras partes interessadas.	X			X		X	X
	1.5. Identificar como a prática de avaliação pode promover a justiça social e o bem público.	X	X	X				
	1.6. Envolver-se em redes e atividades profissionais e contribuir para a profissão de avaliação e sua comunidade de prática.			X	X		X	
	1.7. Utilizar evidências de forma sistemática para fazer julgamentos avaliativos.		X				X	X
	1.8. Fornecer uma perspectiva independente e equilibrada em todos os aspectos da avaliação.			X			X	X
	1.9. Demonstrar flexibilidade.				X			X
	1.10. Apreciar as ligações entre a avaliação e as ciências sociais.						X	X
	1.11. Defender o campo da avaliação e seu valor.		X					
	1.12. Estar comprometido com a transparência em todos os aspectos da avaliação.			X				
	1.13. Demonstrar credibilidade profissional, discrição e confidencialidade ao longo dos processos de avaliação.				X			
	1.14. Praticar dentro da própria competência.				X			
	1.15. Praticar meta-avaliação, buscando formas de incorporar a responsabilidade em seu trabalho.				X			
	1.16. Entender e procurar criar autoeficácia.				X			
	1.17. Oferecer <i>feedback</i> formativo para a equipe e colegas de avaliação.				X			
	1.18. Entender que, mesmo em grupos com a mesma aparência da equipe de avaliação, haverá aspectos culturais diferentes dos próprios avaliadores.				X			
	1.19. Ser sensível, respeitar e cumprir os protocolos e práticas culturais da avaliação.				X			
	1.20. Mostrar sensibilidade ética em contextos sociopolíticos específicos.							

(continua)

(continuação)

DIMENSÃO	COMPETÊNCIAS	KING	AEA	CES	AES	ANZEA	EES	UKES
Metodologia/ Investigação Sistemática	2.1. Determinar e justificar métodos apropriados (quanti, quali e mistos) para responder às questões de avaliação, maximizar o aprendizado e evitar preconceitos.	X	X	X	X	X	X	X
	2.2. Coletar, analisar, interpretar e sintetizar os dados da avaliação usando métodos e critérios apropriados.	X	X	X	X	X	X	X
	2.3. Desenvolver um projeto de avaliação coerente com os objetivos da avaliação.	X		X	X	X	X	X
	2.4. Usar evidências e interpretações para tirar conclusões, fazendo julgamentos e recomendações, quando apropriado.	X	X	X		X	X	X
	2.5. Compreender a base de conhecimento da avaliação (termos, conceitos, teorias, suposições) e revisar a literatura.	X	X		X		X	X
	2.6. Identificar os pressupostos e usar a lógica, metodologia e teorias do programa de avaliação apropriadamente.	X	X	X			X	X
	2.7. Discriminar tópicos e questões da avaliação.	X	X	X			X	X
	2.8. Identificar fontes relevantes de evidências e procedimentos de amostragem, incluindo o escopo e a validade de cada fonte.	X	X	X	X			X
	2.9. Avaliar a confiabilidade e validade dos dados, por meio do uso de verificações de dados, ensaios de controle e comparação, triangulação de resultados e análises cruzadas.	X	X		X			X
	2.10. Identificar propósitos e necessidades de avaliação, justificando as decisões ao longo da avaliação.	X	X	X		X		
	2.11. Relatar de vários ângulos se as constatações da avaliação são credíveis, úteis e acionáveis, se respondem às perguntas feitas e se têm clareza e transparência quanto à metodologia, escolhas e interpretações avaliativas feitas.	X		X		X	X	
	2.12. Integrar sistematicamente evidências de várias fontes, identificando ligações causais, gerando descobertas válidas e observando pontos fortes, lacunas de dados e limitações nas avaliações.	X			X			
	2.13. Realizar meta-avaliações.	X						
	2.14. Avaliar a avaliabilidade do programa.			X				
	2.15. Realizar avaliações de impacto com base na lógica e contexto específicos do programa/projeto de avaliação, verificando os impactos reais e percebidos e a probabilidade de os resultados serem mantidos.				X			
	2.16. Fornecer pensamento crítico e habilidades analíticas e de síntese para a avaliação.					X		
	2.17. Adaptar desenhos e métodos de avaliação a contextos específicos.						X	
	2.18. Utilizar métodos experimentais e quase experimentais nas avaliações.						X	
	2.19. Realizar estudos de caso, pesquisas, entrevistas, painéis de especialistas para aprimorar as avaliações.						X	
	2.20. Compreender abordagens participativas e os contextos nos quais são mais apropriadas.							X
	2.21. Perceber quais métodos são apropriados para fins formativos /somativos.							X

(continua)

(continuação)

DIMENSÃO	COMPETÊNCIAS	KING	AEA	CES	AES	ANZEA	EES	UKES
Análise Situacional/ Contexto	3.1. Identificar as necessidades das partes interessadas e sua capacidade de participar, reconhecendo, respeitando, respondendo aos aspectos da diversidade e mantendo relacionamentos confiáveis.	X	X	X	X	X	X	X
	3.2. Reunir a análise contextual e o engajamento para que a evidência, a análise, a síntese e a interpretação da atividade avaliativa sejam credíveis e válidas para o leque de pessoas envolvidas e afetadas pela avaliação.	X	X	X	X	X	X	X
	3.3. Comunicar processos de avaliação e resultados de maneira oportuna, apropriada e eficaz aos usuários pretendidos.	X	X	X	X			X
	3.4. Identificar, entender, articular e levar em conta o contexto e a situação mais amplos e relevantes para a avaliação.	X	X	X		X	X	X
	3.5. Descrever o programa de avaliação, incluindo sua finalidade básica, componentes e seu funcionamento em contextos mais amplos.	X	X				X	X
	3.6. Promover o uso da avaliação e sua influência no contexto da avaliação.	X	X		X			X
	3.7. Analisar os aspectos políticos relevantes para a avaliação.	X			X			X
	3.8. Resolver conflitos.	X					X	X
	3.9. Determinar a avaliabilidade do programa de avaliação.	X			X			
	3.10. Monitorar e responder a mudanças organizacionais e mudanças no ambiente do programa durante o curso da avaliação.	X		X				
	3.11. Respeitar a singularidade do local de avaliação e do cliente.	X	X					
	3.12. Modificar a avaliação sempre que necessário.	X						X
	3.13. Utilizar o entendimento contextual em análise, síntese, interpretação avaliativa e elaboração de relatórios.				X			X
	3.14. Formar uma equipe de avaliação que tenha credibilidade nesse contexto e alcance de conexões/ relacionamentos, conhecimentos, habilidades e experiência relevantes para a avaliação.					X		
	3.15. Identificar e incorporar protocolos culturais apropriados para interagir com a comunidade, incluindo a incorporação de conhecimentos culturais na equipe de avaliação.				X			
	3.16. Aplicar padrões de maneira sensível ao(s) contexto(s) cultural(is).				X			
	3.17. Usar métodos culturalmente adequados para consulta, engajamento, processos de avaliação e elaboração de relatórios.				X			
	3.18. Rejeitar processos de avaliação que não respeitem normas culturais.				X			
	3.19. Entender e articular o potencial e as limitações da avaliação no(s) contexto(s) cultural(is).				X			
	3.20. Identificar barreiras e facilitadores que influenciarão o esforço de avaliação.				X			
	3.21. Articular como essas relações de poder podem influenciar processos ou resultados específicos de avaliação.				X			

(continua)

(continuação)

DIMENSÃO	COMPETÊNCIAS	KING	AEA	CES	AES	ANZEA	EES	UKES
Gerenciamento de Projeto de Avaliação	4.1. Definir, negociar e gerenciar um projeto de avaliação.	X	X	X	X	X	X	X
	4.2. Identificar os recursos necessários para a avaliação, como informações, conhecimento, pessoal, instrumentos e financeiros.	X	X	X	X		X	X
	4.3. Supervisionar as outras partes interessadas na condução da avaliação.	X	X	X	X		X	X
	4.4. Selecionar e utilizar efetivamente ferramentas de avaliação apropriadas e adequadas à finalidade.	X	X	X	X			X
	4.5. Coordenar e supervisionar processos e produtos de avaliação.		X		X		X	X
	4.6. Definir parâmetros de trabalho, planos e acordos formais para a avaliação.	X		X	X			
	4.7. Utilizar habilidades de gerenciamento e facilitação de equipes de avaliação.		X	X	X			
	4.8. Comunicar-se com os clientes durante todo o processo de avaliação.	X		X				
	4.9. Documentar processos e produtos de avaliação.		X		X			
	4.10. Planejar o uso e influência da avaliação.		X					X
	4.11. Converter conhecimentos e teorias da avaliação em um plano viável para conduzir a avaliação.				X		X	
	4.12. Conduzir a avaliação de maneira ininterrupta.	X						
	4.13. Terminar o trabalho em tempo hábil.	X						
	4.14. Responder às propostas de avaliação.	X						
	4.15. Fazer o orçamento e justificar o custo da avaliação, considerando as necessidades de informação.	X						
	4.16. Treinar outros envolvidos na condução da avaliação.	X						
	4.17. Gerenciar e proteger os dados de avaliação.		X					
	4.18. Demonstrar a capacidade de inovar.				X			
	4.19. Identificar e mitigar problemas/questões de maneiras construtivas e úteis.				X			
	4.20. Abordar aspectos da cultura no planejamento e gerenciamento de avaliações.		X					
	4.21. Relatar ao público de forma justa e no interesse público o que foi acordado.							X
Prática Reflexiva	5.1. Ter consciência de si mesmo como avaliador (conhecimentos, habilidades, atitudes).	X	X	X	X	X	X	X
	5.2. Refletir sobre a prática de avaliação (competências e áreas pessoais de crescimento).	X	X	X	X	X	X	X
	5.3. Prosseguir o desenvolvimento profissional na avaliação.	X	X	X	X	X	X	X
	5.4. Construir relacionamentos profissionais para aprimorar a prática de avaliação.	X	X				X	X
	5.5. Contribuir para a profissão de avaliação.					X	X	X
	5.6. Prosseguir o desenvolvimento profissional em áreas de conteúdo relevante.	X						X
	5.7. Reconhecer e ser transparentes sobre sua influência no processo de avaliação.					X		X

(continua)

(continuação)

DIMENSÃO	COMPETÊNCIAS	KING	AEA	CES	AES	ANZEA	EES	UKES
Competência Interpessoal	6.1. Facilitar a interação interpessoal construtiva (trabalho em equipe, facilitação de grupo, processamento).	X	X	X	X	X	X	X
	6.2. Demonstrar habilidades na comunicação verbal e não verbal eficaz, apropriada e respeitosa.	X	X	X	X		X	X
	6.3. Utilizar habilidades de resolução de conflitos.	X	X	X	X		X	X
	6.4. Colaborar e formar parcerias com as partes interessadas para envolvê-las nos processos de avaliação.		X	X	X	X	X	X
	6.5. Usar habilidades e tecnologias de comunicação escrita na prática de avaliação.	X	X	X	X			X
	6.6. Usar habilidades de negociação.	X			X		X	X
	6.7. Demonstrar competência transcultural.	X	X	X	X			
	6.8. Criar confiança ao longo da avaliação.		X		X			
	6.9. Escutar, entender e engajar diferentes perspectivas.		X					X
	6.10. Manter uma perspectiva objetiva sobre avaliação.				X			X
	6.11. Demonstrar consciência de gênero e sensibilidade cultural.						X	X

Fonte: Elaboração dos autores a partir de King et al. (2001), AEA (2018), CES (2018), AES (2013), Anzea (2011), EES (2011) e UKES (2012).