

INSTRUMENTO PARA DIAGNÓSTICO DE LACUNAS DE DESEMPENHO CONTEXTUAL NO SETOR PÚBLICO

INSTRUMENT FOR DIAGNOSING CONTEXTUAL
PERFORMANCE GAPS IN THE PUBLIC SECTOR

LAÍS CAMPOS DE CARVALHO RODRIGUES

laiscampos123@gmail.com

Universidade de Brasília

<https://orcid.org/0009-0008-3538-8347>

BRUNA STAMM DE BARROS BARRETO

brunastamm99@gmail.com

Universidade de Brasília

<https://orcid.org/0009-0003-2683-9664>

MARIANA LOPES DE ARAÚJO

marylopesaraujo@gmail.com

Universidade de Brasília

<https://orcid.org/0000-0001-6619-3093>

PEDRO PAULO MURCE MENEZES

pemeneses@unb.br

Universidade de Brasília

<https://orcid.org/0000-0002-6937-3302>

FLAVIO GARCIA PRACIANO

flavio.praciano@redes.unb.br

Universidade de Brasília

<https://orcid.org/0000-0002-9453-397X>

RESUMO

Objetivo: Desenvolver e validar um instrumento para diagnóstico de lacunas de desempenho contextual em um órgão público de natureza jurídica, contribuindo para o aprimoramento das práticas de gestão de desempenho no setor público.

Proposta: O estudo propõe um instrumento de avaliação de desempenho construído a partir da realidade institucional, combinando rigor psicométrico e legitimidade organizacional, como alternativa às medidas genéricas frequentemente criticadas por sua baixa aderência ao contexto público.

Abordagem teórica: A pesquisa fundamenta-se na literatura de gestão de desempenho e administração pública, adotando a concepção de desempenho como construto comportamental e distinguindo o desempenho contextual como dimensão transversal às funções exercidas pelos servidores.

Provocação: O artigo provoca o debate sobre a necessidade de instrumentos de avaliação de desempenho sensíveis às especificidades institucionais do serviço público, capazes de ampliar a legitimidade do processo avaliativo e apoiar decisões baseadas em evidências na gestão de pessoas.

Métodos: Adotou-se uma abordagem metodológica mista, integrando etapas qualitativas e quantitativas. Foram realizadas entrevistas com profissionais da carreira jurídica, seguidas da elaboração e validação semântica de descritores, validação por juízes e aplicação do questionário. A validação psicométrica incluiu análises fatoriais exploratória e confirmatória.

Resultados: Os resultados indicaram uma estrutura multifatorial consistente, com evidências satisfatórias de confiabilidade, validade convergente e validade discriminante, confirmando a robustez do instrumento desenvolvido para mensuração do desempenho contextual.

Conclusões: O instrumento validado apresenta contribuição científica e prática ao oferecer uma medida empiricamente robusta e institucionalmente legitimada, capaz de subsidiar políticas de desenvolvimento, capacitação e gestão da força de trabalho no setor público, fortalecendo a efetividade das políticas públicas.

Palavras-chave: Gestão de desempenho; Administração pública; Desempenho contextual; Validação de instrumento; Psicometria.

ABSTRACT

Objective: *To develop and validate an instrument for diagnosing contextual performance gaps in a public organization of a legal nature, contributing to the improvement of performance management practices in the public sector.*

Proposal: *The study proposes a performance evaluation instrument grounded in the institutional reality of public organizations, combining psychometric rigor and organizational legitimacy, as an alternative to generic measures often criticized for their limited contextual fit.*

Theoretical approach: *The research is grounded in the performance management and public administration literature, adopting a behavioral conception of performance and conceptualizing contextual performance as a transversal dimension across public service roles.*

Provocation: *The article seeks to foster debate on the need for performance evaluation instruments that are sensitive to public sector specificities, capable of enhancing evaluative legitimacy and supporting evidence-based people management decisions.*

Methods: *A mixed-methods approach was adopted, integrating qualitative and quantitative stages. Interviews were conducted with legal career professionals, followed by the development and semantic validation of descriptors, expert judgment validation, and nationwide survey application. Psychometric validation included exploratory and confirmatory factor analyses.*

Results: *The findings revealed a consistent multifactorial structure, with satisfactory evidence of reliability, convergent validity, and discriminant validity, confirming the robustness of the instrument for measuring contextual performance.*

Conclusions: *The validated instrument offers scientific and practical contributions by providing an empirically robust and institutionally legitimized measure, supporting public workforce development, training, and performance management policies, and enhancing the effectiveness of public policies.*

Keywords: *Performance management; Public administration; Contextual performance; Instrument validation; Psychometrics.*

1 INTRODUÇÃO

A gestão de desempenho vem assumindo papel cada vez mais relevante como ferramenta estratégica na Administração Pública, ao favorecer o alinhamento entre práticas institucionais e objetivos organizacionais (Lacerda, Ensslin & Ensslin, 2012; Martins & Ensslin, 2021). Nos últimos anos, o fortalecimento de uma cultura orientada para resultados tornou-se um eixo central das políticas de gestão de pessoas nos governos dos países desenvolvidos e do Brasil (OECD, 2021; OECD, 2022). Esse movimento impulsionou a evolução da gestão de pessoas por meio de diferentes fases e marcos regulatórios, voltados ao reforço dos princípios da Gestão Estratégica de Pessoas (GEP) e à valorização da excelência na atuação dos servidores (Cascio, 2015). Entre esses marcos, destaca-se a Política Nacional de Desenvolvimento de Pessoal (PNDP), instituída pelo Decreto nº 5.707/2006 e atualizada pelo Decreto nº 9.991/2019, que orienta o desenvolvimento contínuo dos servidores com base na gestão por competências e na adoção de planos gerais e individuais de desenvolvimento. Outro avanço importante foi a criação do Programa de Gestão e Desempenho (PGD), instituído em 2020 pela Instrução Normativa nº 65 e aperfeiçoado pela Instrução Normativa nº 24/2023, que define orientações, critérios e procedimentos para a pactuação de entregas e o monitoramento de resultados.

Neste sentido, a adoção de instrumentos de avaliação de desempenho (AD) no setor público exige mais do que a mera replicação ou adaptação de modelos consagrados no setor privado (Martins & Ensslin, 2021). A eficácia desses instrumentos está intrinsecamente vinculada à sua capacidade de refletir as especificidades do serviço público, um contexto marcado por finalidades não lucrativas, forte normatização, estabilidade de vínculos, multiplicidade de stakeholders e elevado grau de *accountability*. Diferenças nas culturas organizacionais, nos processos decisórios, nos arranjos institucionais e nas práticas de trabalho tornam inadequado o uso de ferramentas descoladas da realidade estatal (Cascio, 2015; McAdam et al., 2005).

Quando os instrumentos de avaliação carecem de aderência ao contexto público, comprometem-se a confiabilidade dos resultados e a legitimidade do processo, gerando resistências entre avaliadores e avaliados (Eusébio, 2023). Além disso, a presença de subjetividades e vieses, como leniência, efeito halo, tendência central e recenticidade, ainda é um desafio recorrente, dificultando diagnósticos consistentes e a tomada de decisão baseada em evidências. Nesse cenário, torna-se indispensável desenvolver instrumentos ancorados em descritores empiricamente validados, construídos de forma colaborativa com os agentes institucionais e sensíveis às nuances das funções desempenhadas pelos servidores (Belle et al., 2017; Bezerra et al., 2022).

Este artigo apresenta o processo metodológico de construção e validação de um instrumento para diagnóstico de lacunas de desempenho em um órgão público de natureza jurídica, concebido especificamente para superar as limitações de aderência e legitimidade observadas em instrumentos genéricos ou importados do setor privado (Martins & Ensslin, 2021). Ao ancorar-se em descritores empiricamente validados, construídos de forma participativa com agentes institucionais e alinhados às especificidades do contexto público, o estudo contribui para o avanço da gestão de desempenho no setor público ao oferecer um instrumento legítimo, sensível à realidade estatal e capaz de produzir diagnósticos confiáveis para subsidiar decisões gerenciais e políticas de desenvolvimento de pessoas. O processo metodológico envolveu a combinação de etapas qualitativas e quantitativas, contemplando análise da literatura, realização de entrevistas, a validação de descritores e a aplicação em âmbito nacional junto a profissionais da carreira jurídica.

2 REFERENCIAL TEÓRICO

A partir da Reforma do Estado nos anos 1990, a Administração Pública brasileira passou a adotar princípios gerenciais, priorizando eficiência, resultados e foco no cidadão (Camargo et al., 2016). Nesse contexto, a gestão de pessoas assumiu papel estratégico ao alinhar competências individuais às necessidades organizacionais, promovendo tanto desempenho quanto valorização dos servidores, em equilíbrio entre a rigidez normativa e a demanda por agilidade (Deodato, 2021; Souza Camões, 2019).

Diante desse cenário, e em consonância com movimentos já consolidados em contextos internacionais voltados ao estudo de competências essenciais para um serviço público profissional, estratégico e inovador (OECD, 2017), a Escola Nacional de Administração Pública (Enap), a partir de uma série de estudos, propôs um conjunto de competências transversais consideradas indispensáveis ao exercício da função pública. Segundo a ENAP (2021), essas competências são definidas como o “conjunto de conhecimentos, habilidades e atitudes indispensáveis ao exercício da função pública, que contribuem para a efetividade dos processos de trabalho em diferentes contextos organizacionais” (pág. 05). A matriz elaborada pela escola indica sete competências transversais, são elas: Visão sistêmica, Orientação por valores éticos, Trabalho em equipe, Comunicação, Mentalidade digital, Foco nos resultados para os cidadãos, Resolução de problemas com base em dados. Essa estrutura constitui uma iniciativa promissora para a construção de um referencial comum de competências, ao mesmo tempo em que preserva a flexibilidade necessária para adaptação aos diferentes contextos institucionais (OECD, 2023).

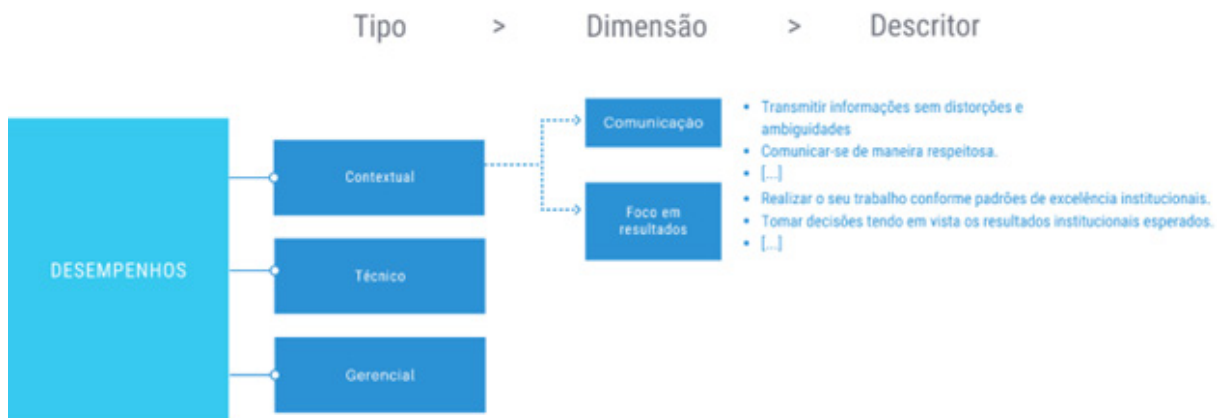
Na literatura, são frequentes as múltiplas classificações do conceito de competências, assim como as diferentes categorias utilizadas para designar aquelas de caráter mais abrangente, aplicáveis a indivíduos em variados contextos, independentemente de suas funções específicas. Tais competências podem ser denominadas transversais, gerais, contextuais, organizacionais, entre outras. Este estudo adota a literatura de gestão de desempenho como referência, dada a heterogeneidade conceitual do termo competência (Nicola & Vosgeral, 2019; Salman et al., 2020) e em função de seu objetivo central: desenvolver um instrumento de diagnóstico de performance individual, voltado a comportamentos expressos, resultados e realizações. Nessa perspectiva, o desempenho é compreendido como comportamento observável, e não apenas como potencial individual derivado de conhecimentos, habilidades e atitudes.

Para fins de mensuração, deve ser concebido como um construto de natureza comportamental que, originado em fatores cognitivos, afetivos, interpessoais e materiais, manifesta-se nas ações intencionais dos indivíduos e se converte, a depender do contexto material e psicossocial de trabalho, em resultados que geram valor econômico e social (Brandão, 2007; Dutra, 2004; Fleury & Fleury, 2001).

Sob a perspectiva conceitual adotada, o desempenho pode ser classificado quanto ao seu tipo, apresentando, na maioria dos casos, três categorias, que podem receber distintas denominações na literatura: gerencial, aplicável aos indivíduos que exercem funções de chefia e liderança; técnico, associado a determinados processos de trabalho e atividades; e contextual, aplicável a todos os indivíduos de determinado contexto independente de suas funções, sendo este último o foco da pesquisa apresentada (Campbell et al., 1993; Sonnentag & Frese, 2002).

Em cada tipo, o desempenho se desdobra em dimensões, que representam grandes temas análogos às competências, compostas por descritores, itens que caracterizam comportamentos esperados e são avaliados por meio de escalas nos formulários de avaliação de desempenho. Em síntese, adota-se um modelo teórico-conceitual que se desdobra em três níveis: tipo de desempenho, dimensão de desempenho e descritores de desempenho. Esses elementos fundamentam a construção das matrizes de desempenho, que estruturam os formulários de avaliação e constituem insumo central das políticas de Gestão de Desempenho (GD) (Aguinis, 2013), conforme ilustrado na Figura 1.

Figura 1. Modelo conceitual dos desempenhos: tipos, dimensões e descritores comportamentais



Fonte: elaborado pelos autores.

A avaliação de desempenho, enquanto componente estruturante do sistema de GD, configura-se como um processo sistemático e contínuo de mensuração da atuação dos servidores, essencial para o alcance dos objetivos das políticas de gestão de pessoas orientadas a resultados. Sua implementação possibilita identificar lacunas de desenvolvimento, por meio de instrumentos avaliativos baseados em escalas de mensuração, e oferece subsídios para decisões estratégicas em gestão de pessoas, especialmente nas frentes de capacitação, progressão funcional e mobilidade na carreira (Bezerra et al., 2022).

Apesar dos avanços observados na literatura e nas práticas institucionais, a mensuração do desempenho contextual no setor público ainda permanece como um desafio. A compreensão desse tipo de desempenho é particularmente relevante, pois permite captar as dinâmicas comportamentais que sustentam a cooperação, a adaptabilidade e a criação de valor público, aspectos que extrapolam as tarefas técnicas e refletem o comprometimento dos servidores com os resultados organizacionais.

A metodologia apresentada na seção seguinte ancora-se em referenciais teóricos consolidados e evidências empíricas da realidade institucional, culminando na proposição de um instrumento de diagnóstico do desempenho contextual. O modelo desenvolvido busca oferecer uma base metodológica robusta, validada e passível de adaptação a diferentes esferas e contextos da Administração Pública, contribuindo para o aperfeiçoamento das práticas de gestão de desempenho e para o fortalecimento de uma cultura organizacional orientada a resultados, colaboração e aprendizagem contínua.

3 MÉTODO

O estudo adotou os pressupostos da pesquisa-ação, entendida como uma abordagem metodológica que articula produção de conhecimento e intervenção prática, por meio da participação ativa dos atores envolvidos no contexto investigado, visando simultaneamente à compreensão do fenômeno e à transformação da realidade estudada (Thiollent, 2011).

Ainda, a pesquisa foi estabelecida em parceria com uma universidade pública e uma organização pública jurídica para construção e validação de um instrumento de avaliação de desempenho profissional. O delineamento do estudo compreendeu etapas qualitativas e quantitativas, organizadas de forma sequencial e complementar (Creswell & Plano Clark, 2011).

A etapa inicial consistiu na coleta de dados qualitativos por meio de entrevistas semiestruturadas com profissionais da carreira do órgão, visando identificar percepções, práticas e expectativas relacionadas ao desempenho no contexto organizacional. Antes da coleta definitiva, realizou-se uma entrevista-piloto para testar o roteiro, aperfeiçoar a abordagem e ajustar a condução das entrevistas. Em seguida, foram conduzidas 12 entrevistas, número considerado suficiente para garantir a saturação teórica das categorias analíticas (Fontanella et al., 2011).

Os dados obtidos foram submetidos à análise de conteúdo, com posterior categorização dos trechos empíricos (Bardin, 2016). A partir dessa categorização, procedeu-se à elaboração de descritores de desempenho fundamentados nos dados empíricos. A partir dos descritores de desempenho elaborados com base na etapa qualitativa, deu-se início ao processo de validação do instrumento. Inicialmente, os descritores foram submetidos a duas etapas preliminares de validação semântica, realizadas antes da validação formal por juízes. A primeira etapa consistiu na validação acadêmica, conduzida por um dos professores coordenadores da pesquisa, que avaliou a pertinência teórica e metodológica dos descritores, realizando ajustes redacionais, refinamentos conceituais e, quando necessário, sugerindo a inclusão de novos itens.

Em seguida, procedeu-se à validação semântica interna no âmbito do órgão, conduzida pela equipe técnica de apoio à pesquisa, composta por servidores da carreira que acompanharam integralmente as etapas de coleta e análise dos dados. Nessa fase, os descritores foram avaliados quanto à clareza, compreensibilidade e adequação terminológica, bem como quanto à sua aderência à realidade organizacional do órgão, incluindo sugestões de ajustes, realocação entre categorias e comentários gerais sobre o instrumento. Essa etapa foi realizada por meio de reuniões técnicas e edições colaborativas dos documentos, assegurando um processo de construção compartilhado e alinhado aos objetivos do estudo.

Concluídas as etapas de validação semântica, os descritores, devidamente ajustados quanto à clareza e pertinência ao contexto organizacional, foram submetidos à validação formal por juízes, com a participação de quatro validadores, etapa que resultou no conjunto final de 34 descritores de desempenho contextual.

Assim, por meio do *Google Forms*, o instrumento foi disponibilizado para aplicação eletrônica, sendo os itens avaliados por meio de uma escala do tipo Likert de cinco pontos de concordância, variando de 1 a 5. A população-alvo da etapa quantitativa foi composta por mais de 2.000 profissionais de carreira do órgão, tendo sido obtidas 202 respostas, sendo 200 válidas no período de novembro a dezembro de 2024. O público participante foi majoritariamente composto por

homens (53,5%). Observa-se que a maior parcela dos respondentes (42%) situava-se na faixa etária entre 35 e 44 anos, e que a maioria da amostra se autodeclarou de cor branca (78,9%).

Procedeu-se ao tratamento dos dados obtidos, de modo a viabilizar as análises estatísticas previstas no estudo. As análises foram conduzidas no *software* JASP e tiveram início com a realização de análises descritivas, seguidas da aplicação de testes inferenciais.

Para fins de evidências de validade, foi realizada Análise Fatorial Exploratória (AFE) e Análise Fatorial Confirmatória (AFC), com o objetivo de avaliar a estrutura fatorial da escala. A AFE foi conduzida a partir de uma matriz de correlações policóricas, adotando-se o método de extração *Robust Diagonally Weighted Least Squares* (RDWLS) (Asparouhov & Muthén, 2010). Considerou-se que o mínimo amostral recomendado pela literatura para a AFE foi atendido, no qual se requer a quantidade de 5 a 10 participantes por item da versão de aplicação da escala (Tabachnick & Fidell, 2019).

4 RESULTADOS E DISCUSSÃO

O processo de desenvolvimento do instrumento teve início com a etapa qualitativa, composta pela análise de entrevistas conduzidas junto a profissionais da carreira jurídica do órgão. A partir da literatura especializada e do exame dos relatos empíricos (ENAP, 2021; Montezano et al., 2022), foi possível estruturar um conjunto inicial de sete dimensões de desempenho contextual, que serviram como guia para a elaboração do roteiro de entrevistas e para a organização preliminar dos descritores. Após a análise das entrevistas, as sete dimensões se mantiveram, sendo que duas sofreram alterações e foram substituídas, resultando em: Adaptabilidade, Aprendizado Contínuo, Comunicação, Foco em Resultado, Organização do Trabalho, Relacionamento Interpessoal e Trabalho em Equipe.

A partir das entrevistas, foram elaborados 49 descritores de desempenho contextual, organizados segundo as sete dimensões previamente definidas. Esses descritores passaram por duas etapas de validação semântica, conduzidas respectivamente pela equipe acadêmica e pela equipe técnica do órgão, o que possibilitou ajustes de redação, exclusões e realocação de itens para assegurar clareza e aderência ao contexto organizacional (Borsa et al., 2012). Em seguida, realizou-se a validação formal por juízes, composta por servidores experientes, utilizando como critério mínimo de concordância o índice de 75%. Após esse processo, a lista inicial foi reduzida para 34 descritores de desempenho, considerados adequados para compor a versão preliminar do questionário.

Logo, seguiu-se para a etapa inicial de validação psicométrica do instrumento, a Análise Fatorial Exploratória (AFE), na qual evidências primárias de validade e confiabilidade são investigadas (Field, 2020). Assim, para verificar a fatorabilidade dos dados, analisou-se o índice *Kaiser-Meyer-Olkin* (KMO), que aponta a adequação e a possibilidade de divisibilidade da amostra em fatores (Kerlinger & Lee, 2008). Para tal, valores próximos a 1 representam maior grau de fatorabilidade, sendo que o índice KMO do instrumento alcançou o valor de 0,92, classificado como meritório segundo a literatura (Hair et al., 2019; Kaiser & Rice, 1974).

Para determinar a quantidade de fatores da escala, foi utilizada a análise paralela, método apontado como o mais adequado em 92% dos casos, em comparação aos critérios do scree plot e do autovalor acima de 1,0, que se mostraram apropriados em apenas 57% e 22% das situações, respectivamente (Laros & Puente-Palacios, 2004). Além de sua maior precisão, a análise paralela apresenta ainda a vantagem de sofrer pouca influência do tamanho amostral (Laros & Puente-Palacios, 2004). Delimita-se a quantidade de fatores no momento a partir do qual o valor estimado supera o valor empírico, de modo que se estabeleceram quatro fatores para a escala (Reise et al., 2000).

Todavia, optou-se por manter a solução teórica de sete fatores, uma vez que esta estrutura foi derivada da etapa qualitativa de construção do instrumento, apresenta coerência conceitual e maior utilidade prática para fins de diagnóstico organizacional. Nesse sentido, considerou-se que a análise exploratória deveria ser complementada pela Análise Fatorial Confirmatória (AFC), de forma a verificar empiricamente a adequação do modelo de sete dimensões, em consonância com recomendações metodológicas que valorizam a primazia do modelo teórico em pesquisas aplicadas (Hair et al., 2019).

O primeiro fator foi denominado Foco em Resultado (FR), por tratar de itens que representam a orientação para a entrega de resultados, comprometimento com metas e capacidade de priorização. O segundo fator correspondeu à Comunicação (COM), relacionado à clareza, precisão e efetividade na transmissão de informações. O terceiro fator, denominado Organização do Trabalho (OT), abrangeu aspectos referentes à gestão de prioridades, planejamento e ordenamento das atividades. O quarto fator foi denominado Relacionamento Interpessoal (RELI), por incluir comportamentos ligados à interação respeitosa, cooperação e gestão construtiva de conflitos. O quinto fator correspondeu ao Trabalho em Equipe (TE), englobando colaboração, compartilhamento de responsabilidades e atuação integrada com colegas. O sexto fator, Aprendizado Contínuo (AC), reuniu itens voltados à busca por atualização, reflexão crítica e incorporação de novos conhecimentos ao trabalho. Por fim, o sétimo fator foi denominado Adaptabilidade (ADAP), referente à capacidade de lidar com mudanças, flexibilidade diante de novos contextos e disposição para ajustar práticas de trabalho.

Em sequência, a AFE foi executada por meio da rotação Oblimin, com fatoração pelo método dos mínimos quadrados ponderados, abordagem empregada dada a presunção da existência de correlações, por tratar-se de um estudo de natureza comportamental (Pasquali, 2010). Desse modo, demonstrou-se a validade interna do instrumento, por meio da qualidade de seus itens, seguindo parâmetros indicados por Hair et al. (2019), de manter somente itens com cargas fatoriais superiores a 0,50. No total, foram retidos 32 itens distribuídos nos sete fatores, após a exclusão de 1 item que apresentou carga fatorial abaixo do ponto de corte. As cargas fatoriais variaram de 0,51 a 0,90, com parte dos itens apresentando valores classificados como muito bons ($\geq 0,63$ e $\leq 0,70$) e outros como excelentes ($\geq 0,71$), enquanto os demais se situaram na faixa aceitável ($\geq 0,50$ e $\leq 0,62$), confirmando a robustez do instrumento (Comrey & Lee, 2013).

Por sua vez, o grau de confiabilidade, conhecido também como consistência interna ou precisão do instrumento, foi estimado por meio do ômega de McDonald (ω). Valores de ω superiores a 0,70 são considerados confiáveis pela literatura, enquanto aqueles acima de 0,80 são muito confiáveis (Dunn et al., 2013). No presente estudo, a escala apresentou $\omega = 0,86$, classificado como muito confiável. As análises por dimensão também revelaram índices satisfatórios: Aprendizado Contínuo (0,84), Adaptabilidade (0,84), Comunicação (0,83), Foco em Resultado (0,83),

Organização do Trabalho (0,83), Relacionamento Interpessoal (0,84) e Trabalho em Equipe (0,84). Esses valores confirmam a consistência interna do instrumento e reforçam sua adequação para mensuração do desempenho contextual.

Não obstante se tratar de um esforço inicial de mensuração, a AFE evidenciou uma estrutura consistente para o instrumento. Os sete fatores extraídos explicam conjuntamente 61,4% da variância total, valor considerado bastante satisfatório para pesquisas na área de ciências sociais aplicadas (Hair et al., 2019).

Observou-se, contudo, que alguns itens apresentaram comportamentos específicos durante a análise. Dois itens originalmente associados ao fator Comunicação (COM2 e COM4) mostraram cargas mais adequadas em outros fatores, tendo sido realocados de acordo com a estrutura sugerida pela AFE. Além disso, um item (COM3) apresentou carga fatorial inferior a 0,50, o que indica contribuição mais limitada ao fator. Ainda assim, optou-se por mantê-lo no instrumento nesta etapa inicial, em razão de sua relevância teórica para o construto e por não comprometer de forma significativa a consistência interna do fator. A Tabela 1 apresenta os índices psicométricos da escala após a etapa de AFE.

Tabela 1 - Índices psicométricos da escala após a AFE

Item	Carga Fatorial							Qualidade
	FR	OT	RELI	TE	COM	AC	ADAP	
FR7	0,83							excelente
FR2	0,71							excelente
FR6	0,69							muito boa
FR4	0,68							muito boa
FR5	0,65							boa
FR3	0,61							boa
FR1	0,53							boa
OT1		0,82						excelente
OT4		0,79						excelente
OT3		0,75						excelente
OT5		0,67						muito boa
OT2		0,58						boa
COM4		0,52						boa
OT6		0,51						boa
RELI6			0,80					excelente
RELI4			0,68					muito boa
RELI5			0,64					muito boa
RELI1			0,63					muito boa
RELI3			0,61					boa
COM2			0,55					boa
RELI2			0,52					boa

Item	Carga Fatorial							Qualidade
	FR	OT	RELI	TE	COM	AC	ADAP	
TE1				0,79				excelente
TE3				0,79				excelente
TE2				0,70				muito boa
COM1					0,74			excelente
COM5					0,63			muito boa
COM3					0,44			razoável
AC2						0,66		muito boa
AC1						0,63		muito boa
AC4						0,52		boa
AC3						0,50		boa
ADAP1							0,78	excelente
ADAP2							0,73	excelente
Confiabilidade (ω)	0,83	0,83	0,84	0,84	0,83	0,84	0,84	
Variância total explicada	61,4%							

Fonte: elaborada pelos autores.

Com base nos resultados obtidos na AFE, avançou-se para a etapa de AFC, conduzida por meio da modelagem de equações estruturais. Para a estimação, empregou-se o método de máxima verossimilhança, amplamente utilizado em estudos psicométricos e considerado robusto mesmo em dados oriundos de escalas do tipo Likert, desde que observadas condições amostrais adequadas (Hair et al., 2019; Kline, 2023).

Nesta etapa, foram analisados os principais índices de ajuste, tanto absolutos quanto incrementais, em conformidade com as recomendações da literatura (Hair et al., 2019; Kline, 2023). Entre os índices absolutos, consideraram-se o qui-quadrado normado (χ^2/gf), o RMSEA (*Root Mean Square Error of Approximation*) e o SRMR (*Standardized Root Mean Square Residual*). Quanto aos incrementais, foram examinados o CFI (*Comparative Fit Index*) e o TLI (*Tucker-Lewis Index*). Como parâmetros de referência, adotaram-se valores de χ^2/gf inferiores a 3,0, RMSEA \leq 0,06, SRMR \leq 0,08, CFI \geq 0,90 e TLI \geq 0,90 (Kline, 2023; Marôco, 2021; Byrne, 2016).

Adicionalmente, conforme as recomendações de Byrne (2016) acerca da parcimônia dos modelos de mensuração, avaliou-se a dimensionalidade do construto por meio da comparação entre os modelos uni e multifatorial. Os resultados evidenciaram que a solução unifatorial apresentou índices aquém dos parâmetros de referência, enquanto o modelo multifatorial demonstrou desempenho superior, com valores satisfatórios em todos os critérios analisados e diferença estatisticamente significativa nos valores de qui-quadrado. Esses achados confirmam que a estrutura com sete fatores representa de maneira mais adequada o desempenho contextual, conferindo maior coerência teórica e utilidade prática para fins de diagnóstico organizacional. Os índices de ajuste de ambos os modelos são apresentados na Tabela 2.

Tabela 2 - Índices de ajuste da AFC

Parâmetro	Referência	Modelo Unifatorial	Modelo Multifatorial
NC(χ^2/DF)	≤ 3.0	4.06	1.73
RMSEA	≤ 0.06	0.12	0.06
SRMR	≤ 0.08	0.10	0.06
CFI	≥ 0.9	0.61	0.91
TLI	≥ 0.9	0.58	0.90

Fonte: elaborada pelos autores.

Na sequência, analisaram-se as cargas fatoriais estimadas para cada item, bem como os índices complementares de confiabilidade e validade do modelo. Conforme o critério de Hair et al. (2019), foram mantidos apenas itens com cargas fatoriais iguais ou superiores a 0,50, o que se confirmou para todos os indicadores, assegurando a adequação do modelo. As cargas variaram de 0,67 a 0,85, sendo classificadas em sua maioria como excelentes ($\geq 0,71$), de acordo com os parâmetros de Comrey e Lee (2013). Esses resultados evidenciam a robustez da estrutura confirmatória obtida e a consistência dos fatores extraídos.

Além das cargas fatoriais, foram examinadas medidas adicionais de confiabilidade e validade convergente. O índice ômega de McDonald (ω) apresentou valores entre 0,81 e 0,91, todos acima do limiar de 0,70 recomendado pela literatura, indicando consistência interna satisfatória dos fatores (Hulin et al., 2001; Ursachi et al., 2015). Do mesmo modo, a variância média extraída (AVE) variou de 0,56 a 0,69, superando o ponto de corte de 0,50 e confirmando a validade convergente do instrumento (Fornell & Larcker, 1981). Esses resultados, aliados às cargas fatoriais elevadas, reforçam a robustez do modelo multifatorial e sua adequação para a mensuração do desempenho contextual. Os valores detalhados encontram-se apresentados na Tabela 3.

Tabela 3 - Índices psicométricos após a AFC

Fator	Item	Estimativa Padronizada	E.P	R.C	p	Qualidade da Carga Fatorial	Ômega de McDonald (ω)	Variância Extraída
FR	FR7	0.83	0.03	31.95	***	excelente	0.91	0.58
	FR6	0.80	0.03	27.47	***	excelente		
	FR4	0.76	0.03	23.06	***	excelente		
	FR2	0.74	0.03	21.22	***	excelente		
	FR5	0.74	0.03	21.06	***	excelente		
	FR1	0.72	0.04	19.26	***	excelente		
	FR3	0.71	0.04	18.77	***	excelente		

Fator	Item	Estimativa Padronizada	E.P	R.C	p	Qualidade da Carga Fatorial	Ômega de McDonald (ω)	Variância Extraída
OT	OT1	0.79	0.03	26.27	***	excelente	0.91	0.57
	OT4	0.79	0.03	25.99	***	excelente		
	OT3	0.78	0.03	24.80	***	excelente		
	OT5	0.77	0.03	23.10	***	excelente		
	OT2	0.72	0.04	19.30	***	excelente		
	OT6	0.72	0.04	18.85	***	excelente		
	COM4	0.72	0.04	19.14	***	excelente		
RELI	RELI2	0.78	0.03	24.87	***	excelente	0.90	0.57
	RELI4	0.78	0.03	25.08	***	excelente		
	RELI6	0.78	0.03	24.44	***	excelente		
	RELI5	0.77	0.03	23.13	***	excelente		
	RELI1	0.74	0.04	20.48	***	excelente		
	RELI3	0.74	0.04	20.57	***	excelente		
	COM2	0.67	0.04	15.61	***	muito boa		
TE	TE1	0.85	0.03	31.52	***	excelente	0.87	0.69
	TE3	0.84	0.03	30.48	***	excelente		
	TE2	0.80	0.03	25.39	***	excelente		
COM	COM3	0.84	0.03	28.55	***	excelente	0.82	0.61
	COM5	0.76	0.04	20.85	***	excelente		
	COM1	0.75	0.04	20.50	***	excelente		
AC	AC3	0.81	0.03	25.79	***	excelente	0.83	0.56
	AC1	0.76	0.04	20.79	***	excelente		
	AC2	0.71	0.04	17.26	***	excelente		
	AC4	0.70	0.04	16.71	***	excelente		
ADAP	ADAP1	0.85	0.04	21.17	***	excelente	0.81	0.68
	ADAP2	0.79	0.04	18.54	***	excelente		

Nota: E.P = Erro Padronizado; R.C = Razão Crítica / Teste T; *** = $p < 0,001$
Fonte: elaborada pelos autores.

Além da validade convergente, foi avaliada também a validade discriminante, com o objetivo de assegurar que cada fator da escala mede um construto distinto e não redundante em relação aos demais. Para tanto, adotou-se o critério de Fornell e Larcker (1981), segundo o qual a variância média extraída de cada fator deve ser superior ao quadrado das correlações entre os fatores. Os resultados confirmaram essa condição em todos os casos, indicando que os sete construtos apresentam identidade própria e não sofrem sobreposição excessiva. Esses resul-

tados, apresentados na Tabela 4, reforçam a evidência de validade discriminante da escala. Complementarmente, foi conduzida a análise da taxa *heterotrait-monotrait* (HTMT), que apresentou valores entre 0,40 e 0,77, todos abaixo do ponto de corte de 0,85 (Henseler et al., 2015), corroborando a distinção entre os construtos, ainda que não esteja representada na tabela por convergir com os achados obtidos pelo critério de Fornell e Larcker.

Tabela 4 - Validade divergente da escala

Fator	FR	OT	RELI	TE	COM	AC	ADAP
FR	0,58 ^a	-	-	-	-	-	-
OT	0,28	0,57 ^a	-	-	-	-	-
RELI	0,22	0,22	0,57 ^a	-	-	-	-
TE	0,14	0,26	0,25	0,69 ^a	-	-	-
COM	0,27	0,35	0,46	0,19	0,61 ^a	-	-
AC	0,36	0,18	0,15	0,25	0,12	0,56 ^a	-
ADAP	0,14	0,18	0,14	0,17	0,12	0,20	0,68 ^a

Fonte: elaborada pelos autores.

De forma integrada, os resultados demonstram que o instrumento apresenta propriedades psicométricas sólidas, conciliando consistência interna, validade convergente e validade discriminante. Essa robustez metodológica é particularmente relevante no contexto do serviço público, em que ferramentas de diagnóstico de desempenho são frequentemente criticadas por baixa aderência ao contexto organizacional (Andrade & Cordeiro, 2022; Vilhena & Martins, 2022). No presente estudo, a manutenção da estrutura de sete fatores reforça a adequação do modelo às especificidades da carreira jurídica, oferecendo não apenas uma medida válida, mas também um recurso prático para subsidiar políticas de desenvolvimento e gestão do desempenho.

Além do rigor psicométrico, destaca-se o processo colaborativo de construção do instrumento, que envolveu desde a análise qualitativa de entrevistas até a validação por juízes e a aplicação nacional. Esse percurso metodológico atende a uma demanda recorrente na Administração Pública por instrumentos que sejam simultaneamente cientificamente robustos e institucionalmente legitimados. Ao incorporar descritores validados junto aos profissionais da carreira jurídica, assegurou-se não apenas a clareza semântica dos itens, mas também sua pertinência em relação às funções desempenhadas no cotidiano do órgão, superando a crítica de que avaliações importadas de outros contextos carecem de aderência ao setor público (Tamada & Kowal Olm Cunha, 2023).

Outro aspecto relevante a ser discutido refere-se às implicações práticas do instrumento para a gestão de pessoas. Ao identificar de forma precisa as lacunas de desempenho contextual, a escala possibilita que o órgão direcione investimentos em capacitação, planejamento da força de trabalho e desenvolvimento de lideranças com base em evidências. Trata-se de um avanço frente a modelos tradicionais de avaliação, que frequentemente se limitam a aferir cumprimento de metas ou resultados imediatos (Araújo et al., 2022; Cunha et al., 2016). No contexto jurídico

analisado, a mensuração do desempenho contextual adquire especial relevância por capturar dimensões como adaptabilidade, comunicação e trabalho em equipe, competências críticas para a efetividade da atuação institucional em um ambiente de alta complexidade normativa e múltiplos stakeholders.

4.1 Contribuições, Implicações, Limitações e Agenda

Os achados deste estudo oferecem contribuições relevantes de natureza científica, metodológica e aplicada para o campo da gestão de desempenho no setor público. Do ponto de vista teórico, o principal avanço reside no aprofundamento da literatura sobre avaliação de desempenho em contextos públicos, ao propor e validar empiricamente um instrumento específico para a mensuração do desempenho contextual. Embora esse tipo de desempenho seja amplamente reconhecido na literatura internacional como componente essencial da efetividade organizacional, sua operacionalização no serviço público brasileiro ainda é incipiente e frequentemente subsumida a modelos genéricos ou excessivamente focados em resultados e tarefas. Ao desenvolver um conjunto de dimensões e descritores ancorados na realidade institucional, o estudo contribui para qualificar a compreensão do desempenho como construto comportamental, sensível às especificidades do trabalho público.

Adicionalmente, a validação psicométrica robusta do instrumento, sustentada por análises fatoriais exploratória e confirmatória, bem como por evidências de confiabilidade, validade convergente e discriminante, reforça a qualidade do instrumento proposto e amplia a confiança em sua utilização em contextos organizacionais reais. Ao articular rigor metodológico e aderência empírica, o estudo responde a críticas recorrentes dirigidas aos sistemas de avaliação de desempenho no setor público, pelo menos do ponto de vista da sua capacidade diagnóstica.

Sob a perspectiva metodológica, o estudo apresenta como contribuição o processo de construção do instrumento baseado em triangulação de métodos e fontes. A integração entre entrevistas qualitativas, validação semântica, avaliação por juízes e análises psicométricas configura um percurso metodológico consistente e replicável, alinhado às boas práticas de desenvolvimento de instrumentos. Além disso, o trabalho constitui um exemplo concreto de cooperação entre academia e órgão público, evidenciando a viabilidade de produzir conhecimento científico aplicado sem prejuízo do rigor teórico-metodológico.

No que se refere às implicações práticas, o instrumento desenvolvido possibilita diagnósticos mais precisos das lacunas de desempenho contextual, oferecendo subsídios diretos para o planejamento de ações de capacitação, desenvolvimento de pessoas e gestão do desempenho. Ao permitir a identificação de dimensões comportamentais críticas, a escala amplia o escopo das avaliações no contexto estudado, identificando comportamentos inerentes aos servidores. Ademais, o caráter participativo por parte dos servidores da casa no processo de construção tende a reforçar a legitimidade do instrumento junto aos servidores, reduzindo resistências e favorecendo sua incorporação às rotinas de gestão.

Em termos de políticas públicas, os resultados do estudo podem apoiar a implementação e o aperfeiçoamento de programas de gestão de desempenho e desenvolvimento de pessoas, como o Programa de Gestão e Desempenho (PGD) e a Política Nacional de Desenvolvimento

de Pessoas (PNDP). Ao fornecer uma medida validada e sensível ao contexto institucional, o instrumento contribui para a adoção de práticas baseadas em evidências na administração pública, fortalecendo a capacidade do Estado de alinhar desempenho individual, desenvolvimento profissional e entrega de valor público.

Apesar das contribuições apresentadas, o estudo possui limitações que devem ser consideradas na interpretação dos resultados. A amostra esteve restrita a profissionais de carreira jurídica de um único órgão público, o que limita a generalização dos achados para outros contextos organizacionais e carreiras. Além disso, não foram testadas evidências de validade nomológica e preditiva, uma vez que o desempenho contextual não foi relacionado a outros construtos ou indicadores de resultados organizacionais. Soma-se a isso o delineamento transversal da pesquisa, que impossibilita a avaliação da estabilidade temporal do instrumento por meio de procedimentos como teste-reteste.

Diante dessas limitações, delinea-se uma agenda de pesquisas futuras promissora. Recomenda-se a aplicação do instrumento em outros órgãos públicos e em carreiras jurídicas diversas, com vistas à ampliação de sua validade externa. Estudos adicionais podem testar a invariância de medida entre diferentes grupos, como gênero, tempo de carreira e regiões, contribuindo para a consolidação do instrumento em nível nacional. Ademais, pesquisas longitudinais são necessárias para avaliar a estabilidade temporal e a validade preditiva da escala. Por fim, sugere-se explorar empiricamente as relações entre desempenho contextual e variáveis do comportamento organizacional positivo, como engajamento, bem-estar, liderança e efetividade institucional, ampliando a compreensão dos mecanismos pelos quais o desempenho contextual contribui para a qualidade da gestão pública.

5 CONCLUSÕES

O artigo apresentou o processo metodológico de desenvolvimento de um instrumento diagnóstico de lacunas de desempenho, aplicado em um órgão público de natureza jurídica. A construção coletiva, ancorada em entrevistas qualitativas e validação por especialistas, associada a análises psicométricas rigorosas, assegurou evidências de validade e confiabilidade da escala. Esse percurso confere legitimidade científica e institucional ao instrumento, fortalecendo sua aplicabilidade no contexto da gestão pública.

O questionário ora apresentado representa um modelo replicável e adaptável a diferentes organizações públicas, constituindo-se em ferramenta estratégica para a identificação de lacunas de desempenho contextual. Ao oferecer subsídios empíricos para o planejamento de ações de capacitação, desenvolvimento de pessoas e gestão do desempenho, o instrumento contribui para a formulação de políticas mais alinhadas às necessidades institucionais e às especificidades do serviço público, promovendo maior efetividade na entrega de valor à sociedade.

REFERÊNCIAS

- Aguinis, H. (2013). *Performance management* (3rd ed.). Upper Saddle River, NJ: Pearson.
- Andrade, M. N. S. d., & Cordeiro, A. T. (2022). Avaliação de desempenho no setor público: A perspectiva de servidores em uma instituição federal. *Revista de Carreiras e Pessoas*, 12(2), 259–279. <https://doi.org/10.23925/recape.v12i2.49959>
- Araújo, M. L., Menezes, P. P. M., & Demo, G. (2022). Challenges in implementing competency-based management in the Brazilian public sector: An integrated model. *Sustainability*, 14(22), 14755. <https://doi.org/10.3390/su142214755>
- Bardin, L. (2016). Análise de Conteúdo. São Paulo: Edições 70. In São Paulo: Edições 70 (p. 279).
- Belle, N., Cantarelli, P., & Belardinelli, P. (2017). *Cognitive biases in performance appraisal: Experimental evidence on anchoring and halo effects with public sector managers and employees*. *Review of Public Personnel Administration*, 37(4), 405–429. <https://doi.org/10.1177/0734371X17704891>
- Bezerra, L. F., Cunha, D. de O. da, Severiano Júnior, E., & Zouain, D. M. (2022). Avaliação de desempenho individual na administração pública: Uma revisão integrativa da literatura. *Revista FSA*, 19(1), 3–23. <https://doi.org/10.12819/2022.19.11>
- Borsa, J. C., Damásio, B. F., & Bandeira, D. R. (2012). Adaptação e validação de instrumentos psicológicos entre culturas: Algumas considerações. *Paidéia (Ribeirão Preto)*, 22(53), 423–432. <https://doi.org/10.1590/s0103-863x2012000300014>
- Brandão, H. P. (2007). *Competências no trabalho: uma análise da produção científica brasileira*. Estudos de Psicologia (Natal), 12(2), 149–158
- Byrne, B. M. (2016). *Structural equation modeling with AMOS*. Routledge. <https://doi.org/10.4324/9781315757421>
- Camargo, E. J. B., Pereira, A. S., Grzybovski, D., Lorenzon, A. L. H., & Beltrame, B. (2016). Gestão de pessoas na administração pública federal: Análise dos modelos adotados no período 1995 a 2010. *Revista Interdisciplinar de Gestão Social*, 5(1).
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In E. Schmitt, W. C. Borman, & Associates (Eds.), *Personnel selection in organizations* (pp. 35–70). San Francisco: Jossey-Bass.
- Cascio, W. F. (2015). Strategic HRM: Too important for an insular approach. *Human Resource Management*, 54(3), 423–426. <https://doi.org/10.1002/hrm.21735>
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press. <https://doi.org/10.4324/9781315827506>
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Cunha, C. M., De Almeida Neto, O. P., & Stackfleth, R. (2016). Principais métodos de avaliação psicométrica da validade de instrumentos de medida. *Revista Brasileira Ciências da Saúde - USCS*, 14(47). <https://doi.org/10.13037/ras.vol14n47.3391>
- Deodato, K. R. (2021). *A gestão da qualidade: Um fator chave para a modernização do setor público – um estudo de caso aplicado a uma instituição de ensino superior pública* (Dissertação de mestrado, Universidade de Coimbra). Faculdade de Direito, Universidade de Coimbra.
- Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412. <https://doi.org/10.1111/bjop.12046>
- Dutra, J. S. (2004). *Competências: Conceitos e instrumentos para a gestão de pessoas na empresa moderna*. Atlas.
- Escola Nacional de Administração Pública. (2021). *Competências transversais de um setor público de alto desempenho*. Brasília: Escola Nacional de Administração Pública. <https://repositorio.enap.gov.br/bitstream/1/5663/4/1_Compete%CC%82ncias%20transversais%20de%20um%20setor%20pu%CC%81blico%20de%20alto%20desempenho_16.12.2021.pdf>. Acesso em: 06 de out. 2025.
- Eusébio, T. (2023). *A avaliação de desempenho na administração pública: a comparação entre Portugal e o Gana*. *Revista Portuguesa de Gestão Contemporânea*, 4(1), 13–40.
- Field, A. (2020). *Descobrimo a estatística usando o SPSS* (5. ed.). Penso.
- Fleury, A., & Fleury, M. T. L. (2001). *Estratégias empresariais e formação de competências: Um quebra-cabeça caleidoscópico da indústria brasileira*. Atlas.

- Fontanella, C. B. J. B., Jose, B., Fontanella, B., Luchesi, B. M., Giovana, M., Saidel, B., Ricas, J., Turato, E. R., & Melo, D. G. (2011). Amostragem em pesquisas qualitativas: proposta de procedimentos para constatar saturação teórica. *Caderno de Saúde Pública*, 27(2), 389–394. <http://www.scielo.br/pdf/csp/v27n2/20.pdf>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39. <https://doi.org/10.2307/3151312>
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning EMEA.
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hulin, C., Netemeyer, R., & Cudeck, R. (2001). Can a Reliability Coefficient Be Too High? *Journal of Consumer Psychology*, 10(1), 55–58.
- Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, 34(1), 111–117. <https://doi.org/10.1177/001316447403400115>
- Kerlinger, F. N., & Lee, H. B. (2008). *Foundations of behavioral research* (5. ed.). Wadsworth Thomson.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). New York: The Guilford Press
- Lacerda, R. T. O., Ensslin, L., & Ensslin, S. R. (2012). Uma análise bibliométrica da literatura sobre estratégia e avaliação de desempenho. *Gestão & Produção*, 19(1), 59–78.
- Laros, J. A., & Puente-Palacios, K. E. (2004). Validação cruzada de uma escala de clima organizacional. *Estudos de Psicologia (Natal)*, 9(1), 113–119. <https://doi.org/10.1590/s1413-294x2004000100013>
- McAdam, R., Hazlett, S., & Casey, C. (2005). Performance management in the UK public sector: Addressing multiple stakeholder complexity. *International Journal of Public Sector Management*, 18(3), 256–273.
- Martins, V. A., & Ensslin, S. R. (2021). Avaliação de desempenho no setor público: Oportunidades de pesquisa com base nas lacunas identificadas em estudos de casos. *Revista de Contabilidade do Mestrado em Ciências Contábeis da UERJ*, 25(3). <https://doi.org/10.12979/rcmccuerj.v25i3.51180>
- Marôco, J. (2021). *Análise de equações estruturais: fundamentos teóricos, software & aplicações* (3.ed.). Pêro Pinheiro: ReportNumber
- Montezano, L., Isidro, A., La Falce, J. L., & Sano, H. (2022). Antecedentes das dimensões da gestão pública inovadora brasileira: Competências de equipes e individuais. *Teoria e Prática em Administração*, 12(2), 145–169. <https://doi.org/10.22478/ufpb.2238-104X.2022v12n2.61415>
- Nicola, R. de M. S., & Vosgeral, D. S. A. R. (2019). Conceitos e enfoques em competências nas pesquisas brasileiras: Uma revisão narrativa. *Revista e-Curriculum*, 17(1), 107–144. <https://doi.org/10.23925/1809-3876.2019v17i1p107-144>
- OECD. (2017). *Skills for a high performing civil service. OECD Public Governance Reviews*. OECD Publishing. <https://doi.org/10.1787/9789264280724-en>
- OECD. (2021). *Public Employment and Management 2021: The Future of the Public Service*. OECD Publishing. <https://doi.org/10.1787/938f0d65-en>
- OECD. (2022). *OECD public service leadership and capability review of Brazil*. <https://www.gov.br/servidor/pt-br/assuntos/laboragov/oecd-public-service-leadership-and-capability-review-of-brazil-2.pdf>
- OECD. (2023). *Public employment and management 2023: Towards a more flexible public service*. OECD Publishing. <https://doi.org/10.1787/5b378e11-en>
- Pasquali, L. (2010). *Instrumentos psicossociais: manual prático de elaboração*. Laboratório de Pesquisa em Avaliação e Medida (LabPAM). Instituto de Psicologia. Brasília: Universidade de Brasília.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12(3), 287–297. <https://doi.org/10.1037/1040-3590.12.3.287>
- Salman, M., Ganie, S. A., & Saleem, I. (2020). The concept of competence: A thematic review and discussion. *European Journal of Training and Development*, 44(6/7), 717–742. <https://doi.org/10.1108/EJTD-10-2019-0171>

Sonnentag, S., & Frese, M. (2002). *Performance concepts and performance theory*. In S. Sonnentag (Ed.), *Psychological management of individual performance* (pp. 3–25). John Wiley & Sons.

Souza Camões, M. R. de. (2019). Inovação em gestão de pessoas no setor público: Reflexão sobre as práticas no contexto brasileiro. *Revista de Políticas Públicas e Gestão Governamental*, 16(1), 9–24.

Tamada, R. C. P., & Kowal Olm Cunha, I. C. (2023). Competências profissionais do técnico administrativo em educação: Evidências de validade do conteúdo. *Revista Gestão Universitária na América Latina - GUAL*, 01–22. <https://doi.org/10.5007/1983-4535.2023.e89677>

Thiollent, M. (2011). *Metodologia da pesquisa-ação* (18ª ed.). São Paulo, SP: Cortez.

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How reliable are measurement scales? External factors with indirect influence on reliability estimators. *Procedia Economics and Finance*, 20, 679–686. [https://doi.org/10.1016/s2212-5671\(15\)00123-9](https://doi.org/10.1016/s2212-5671(15)00123-9)

Vilhena, R., & Martins, H. . (2022). Percepções sobre gestão do desempenho no setor público. *Revista Do Serviço Público*, 73(a), 8-25.