

### Samuel Souza de Araújo

Universidade Federal de Santa Catarina (UFSC)

Florianópolis, SC, Brasil

samuel.souza@ufsc.br

ORCID: <http://orcid.org/0000-0002-0446-7578>

### Lucas dos Santos Matos

Universidade Federal de Santa Catarina (UFSC)

Florianópolis, SC, Brasil

L.matos@ufsc.br

ORCID: <http://orcid.org/0000-0001-6736-7227>

### Sandra Rolim Ensslin

Universidade Federal de Santa Catarina (UFSC)

Florianópolis, SC, Brasil

senssln@gmail.com

ORCID: <http://orcid.org/0000-0001-7420-8507>

## Compras Públicas Sob a Perspectiva da Avaliação de Desempenho: Uma Revisão de Literatura e Agenda de Pesquisa

### Public Procurement under the Performance Evaluation Perspective: a Literature Review and Research Agenda

---

#### RESUMO

Esta pesquisa analisou a literatura referente compras públicas, sob a perspectiva da avaliação de desempenho, visando gerar conhecimento acerca do que tem sido realizado e lacunas que demandam investigação para propor uma agenda para avanços nessa área. Utilizou-se o instrumento *Knowledge Development Process-Constructivist* para conduzir o processo de seleção, reflexão e análise dos artigos que compõem o portfólio bibliográfico representativo do tema, por abordarem o fragmento da literatura delimitado. Com base na análise bibliométrica da avaliação de desempenho das compras públicas evidenciou-se: inexistência de autores prolíficos e de rede de colaboração; *stakeholders* como foco de atenção dos pesquisadores; modelos incapazes de oferecer subsídios à gestão das compras públicas; avaliação de desempenho visando avaliar, controlar e motivar compras públicas; inexistência de falhas métricas de replicação de uma instituição para outra e foco excessivo em medidas financeiras. Com base nos achados, propõem-se uma agenda de pesquisa no âmbito acadêmico e gerencial.

**Palavras-Chave:** compras públicas; avaliação de desempenho; revisão de literatura; ProKnow-C.

---

#### ABSTRACT

This research analyzed the literature regarding Public Procurement, from the perspective of performance evaluation, aiming to generate knowledge about what has been accomplished and gaps that demand investigation to propose an agenda for advances in this area. The Knowledge Development Process-Constructivist was applied to conduct the selection, reflection and analysis of the articles that composing the bibliographic portfolio by addressing the delimited literature fragment. Based on the bibliometric analysis of public procurement performance evaluation, it was evidenced: no prolific authors and collaborative network; stakeholders as the focus of researchers' attention; models unable to offer subsidies to the management of public purchases; performance evaluation aiming to evaluate, control and motivate public procurement; inexistence of metric failures of replication from one institution to another and excessive focus on financial measures. Based on the findings, an academic and managerial agenda is proposed.

**Keywords:** public procurement; performance evaluation; literature review; ProKnow-C.

Universidade Federal do Espírito Santo - UFES

Endereço

Av. Fernando Ferrari, 514, Goiabeiras

29.075-910, Vitória-ES

gestaoeconexoes@gmail.com

gestaoeconexoes@ccje.ufes.br

<http://www.periodicos.ufes.br/ppgadm>

Coordenação

Programa de Pós-Graduação em

Administração (PPGADM/CCJE/UFES)

Artigo

Recebido em: 05/09/2019

Aceito em: 16/12/2019

Publicado em: 24/01/2020

## Introdução

O Departamento de Compras Públicas é responsável por adquirir bens, obras e serviços para suprir os órgãos públicos e, conseqüentemente, atender às necessidades dos cidadãos (Mardale, 2015; Patrucco, Luzzini & Ronchi, 2016; Srivastava & Agrahari, 2017; Tillema, Mimba & Helden, 2010). Por vezes, tal departamento acaba pagando um preço mais alto por produtos ou serviços cuja qualidade não se justifica, visto que é possível utilizar mecanismos legais para reduzir os custos e garantir a qualidade desejada, equacionando da melhor forma o valor percebido da qualidade pelo recurso gasto quando da aquisição/pagamento do bem/serviço (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Rotchanakitumnui, 2013; Schiele, 2005; Srivastava & Agrahari, 2017). Tal atividade evidencia a importância do Setor/Departamento de Compras para a realização e a prestação dos serviços do setor público (Patrucco *et al.*, 2016; Rendon, 2008; Schiele, 2005).

Por conta disso, surge a necessidade de avaliar o desempenho dos processos de aquisição (Gardenal, 2013; Patrucco *et al.*, 2016; Rendon, 2008). Essa necessidade tornou-se mais evidente especialmente em função das reformas administrativas influenciadas pela mudança de paradigma introduzidas pela Nova Gestão Pública (*New Public Management* - NPM) em busca da eficiência e da eficácia das organizações públicas (Galera, Rodríguez & Hernandez, 2008; Patrucco *et al.*, 2016; Rendon, 2008). Além disso, a NPM visa a reduzir os custos, melhorar o desempenho do setor público e a qualidade e produtividade dos serviços prestados, apesar das críticas quanto ao excesso de confiança na adoção de técnicas derivadas de empresas privadas (Chomchaiya & Esichaikul, 2016; Galera *et al.*, 2008).

Por meio das mudanças que ocorreram no setor público sob o paradigma da NPM, surge o estímulo ao uso de práticas de avaliação de desempenho (AD) em organizações públicas (Patrucco *et al.*, 2016; Tillema *et al.*, 2010). O alinhamento às práticas de avaliação de desempenho pode ser identificado a partir da definição proposta por Neely, Gregory e Platts (1995, p. 80), pesquisadores renomados nessa área de conhecimento, na qual a AD é tratada como o “processo de quantificação da eficiência e eficácia da ação”. Nesse sentido, evidencia-se a oportunidade de uso da AD para a gestão de compras públicas, de modo a suprir a demanda e a necessidade existentes, promovendo melhorias para o tema. Contudo, o conhecimento acerca do tema encontra-se disperso e ainda pouco explorado, carecendo de estudos que construam conhecimento consolidado referente à relação da AD com as compras públicas, para permitir a compreensão do alinhamento da literatura com as práticas organizacionais no setor.

Nesse contexto, investigar o que tem sido publicado na literatura referente à avaliação de desempenho das compras públicas pode contribuir para subsidiar ações futuras de gestores dos Departamentos de Compras, como também para se propor reflexões e propostas teóricas que conduzam os processos de compras públicas a serem mais eficientes e efetivos. Assim, o objetivo da pesquisa é analisar um fragmento relevante da literatura referente às compras públicas sob a perspectiva da avaliação de desempenho, visando a gerar conhecimento acerca do que tem sido

realizado, quais lacunas demandam investigação, para, então, propor alternativas para avanço nessa área. Para a condução da seleção, análise e reflexão da literatura, o instrumento *Knowledge Development Process-Constructivist (ProKnow-C)* foi escolhido (Matos, S. R. Ensslin & L. Ensslin, 2019; S. R. Ensslin, L. Ensslin, Matos, Dutra & Ripoll-Feliu, 2015; Thiel, S. R. Ensslin & L. Ensslin, 2017; Valmorbida, S. P. R. Ensslin, L. Ensslin & Ripoll-Feliu, 2016).

O estudo se justifica por contribuir para diminuir a carência de estudos que investiguem compras públicas e avaliação de desempenho, conforme apontado por Patrucco *et al.* (2016). Patrucco *et al.* (2016, p. 3) afirmam que há “uma lacuna entre as ações pretendidas e os resultados reais”, quando as reformas influenciadas pela NPM são postas em prática. Surge, assim, a necessidade de medir e melhorar os processos de compras (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Rendon, 2008; Srivastava & Agrahari, 2017). A delimitação deste estudo está na seleção de artigos em língua inglesa, publicados a partir de 2000 nas bases de dados *Scopus*, *Web of Science*, *Science Direct*, *Proquest* e *Emerald*, integrantes do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Após esse capítulo introdutório, será apresentado o referencial teórico do estudo, dividido em Avaliação de Desempenho e Avaliação de Desempenho de Compras Públicas. No terceiro capítulo será apresentada a metodologia da pesquisa e explicado o uso do instrumento de intervenção *ProKnow-C*. Os resultados serão apresentados no quarto capítulo e, na sequência, serão feitas considerações finais.

## Fundamentação teórica

### Avaliação de desempenho

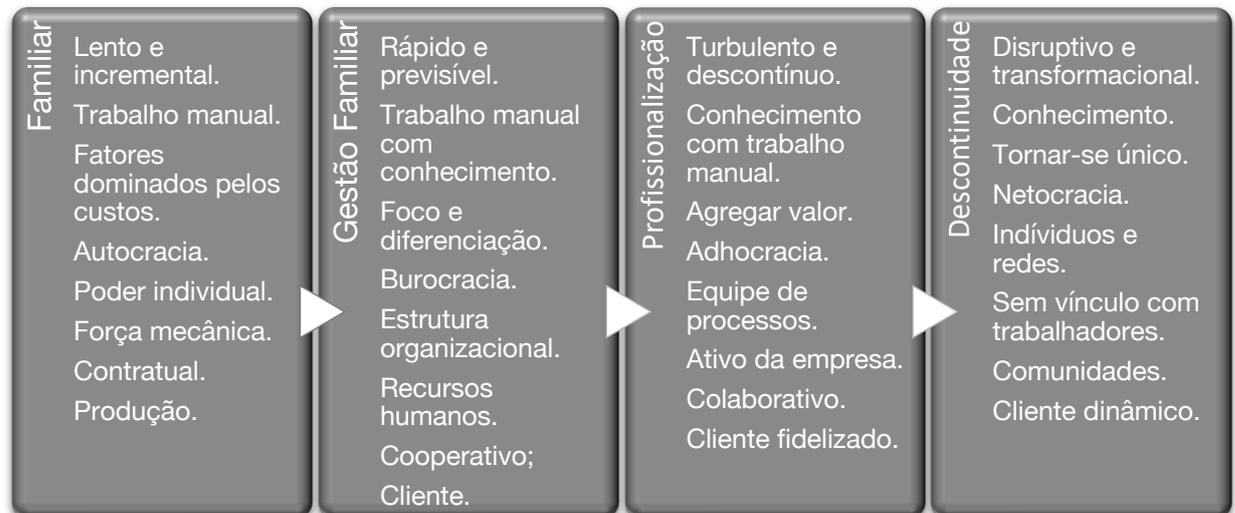
Para melhor compreender o conceito de avaliação de desempenho (AD), buscase inicialmente identificar como ela se desenvolveu ao longo do tempo, para depois definir os conceitos que foram utilizados no estudo. É possível observar uma falta de consenso nas definições de AD (Camp & Braet, 2016; Franco-Santos *et al.*, 2007; Lebas, 1995; Matos *et al.*, 2019; Neely *et al.*, 1995). Apesar da falta de definição, os autores concordam que, na década de 1980, ocorreram as principais alterações na área de avaliação de desempenho devido à insatisfação com as medidas de desempenho tradicionais (Bititci, Garengo, Dörfler & Nudurupati, 2012; Bourne, Mills, Wilcox, Neely & Platts, 2000; Carneiro-da-Cunha, Hourneaux & Corrêa, 2016; Choong, 2014; Franco-Santos *et al.*, 2007; Ghalayini & Noble, 1996; Matos *et al.*, 2019; Melnyk, Bititci, Platts, Tobias & Andersen, 2014; Neely, 1999; Neely *et al.*, 1995; Nudurupati, Bititci, Kumar & Chan, 2011). A evolução das medidas é apresentada na Tabela 1.

Tabela 1  
Evolução das Medidas de Desempenho

Medidas de Desempenho Tradicionais	Medidas de Desempenho Não Tradicionais
Baseado no Sistema de Contabilidade	Baseado na estratégia da empresa
Principalmente medidas financeiras	Principalmente medidas não financeiras
Destinado aos gerentes	Destinado a todos os funcionários
Métricas desatualizadas (foco no passado)	Métricas atuais (presente)
Díficeis, confusas e enganosas	Simples, preciso e fácil de usar
Leva à frustração dos funcionários	Conduzir à satisfação dos funcionários
Negligenciado no chão da fábrica	Frequentemente usadas no chão de fábrica
Fixo (Formato, locais e durante o tempo)	Variável conforme a necessidade
Monitorar (controle) do desempenho	Melhorar o desempenho (melhoria contínua)

**Nota.** Adaptado de "The changing basis of performance measurement" (p. 68). Fonte: Ghalayini, A. M., & Noble, J. S. (1996). *International Journal of Operations & Production Management*, 16(8), 63-80.

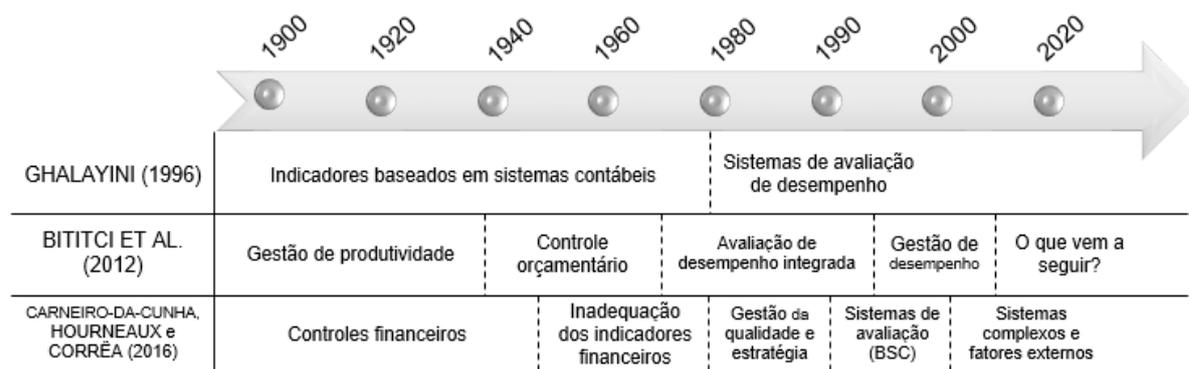
A evolução das medidas de desempenho acompanharam as mudanças do ambiente de negócios (Bititci *et al.*, 2012; Bourne *et al.*, 2000; Neely, 1999). Em seu estudo, Bititci *et al.* (2012) afirmam que tais evoluções ocorreram em resposta às tendências globais e que é possível traçar um paralelo com a evolução dos contextos empresariais, conforme mostra a Figura 1.



**Figura 1.** Contextos Empresariais

Fonte: Adaptado de "Performance measurement: challenges for tomorrow" (p. 312), Bititci *et al.* (2012). *International Journal of Operations & Production Management*, 14(3), 305-327.

Com base nos contextos e nas características específicas da organização, pode-se determinar, inclusive, a fase na qual se encontra a avaliação de desempenho. Na Figura 2 estão as fases evolutivas da AD, conforme a concepção de Ghalayini e Noble (1996), Bititci *et al.* (2012) e Carneiro-da-Cunha *et al.* (2016).



**Figura 2.** Evolução da Avaliação de Desempenho

Ghalayini e Noble (1996) dividem a avaliação de desempenho em duas fases: a primeira com ênfase em medidas financeiras; e a segunda com ênfase no desenvolvimento de sistemas de avaliação de desempenho (SAD). Bititci *et al.* (2012) aprofundam isso, afirmando que gestão da produtividade, controle orçamentário, avaliação de desempenho integrado e gestão de desempenho integrado são as fases evolutivas do campo da AD.

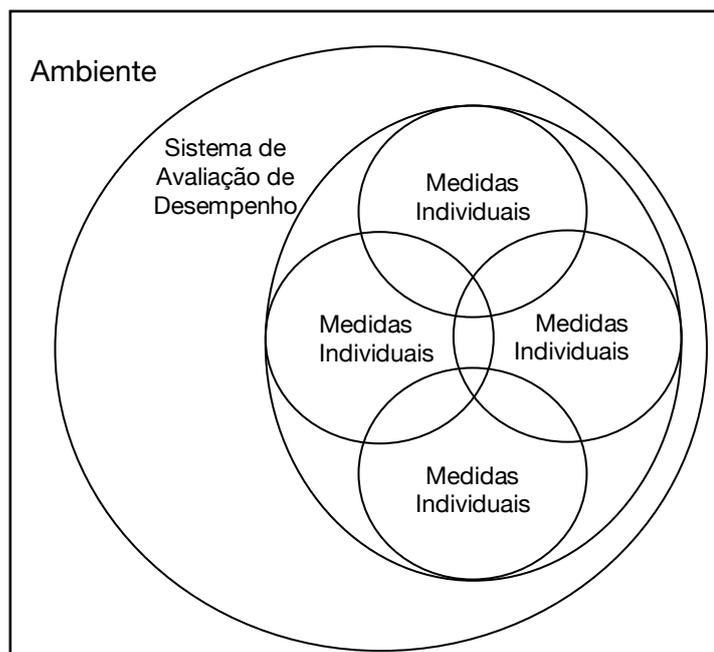
Com a evolução histórica, é possível compreender o desenvolvimento do tema ao longo do tempo para então definir avaliação de desempenho como o “processo de quantificação da eficiência e eficácia da ação” (Neely *et al.*, 1995, p. 80). A métrica é utilizada para essa quantificação, definida como medida de desempenho, enquanto o sistema de avaliação de desempenho é o conjunto dessas métricas (Neely *et al.*, 1995).

Melnyk *et al.* (2014) aprofundam o conceito de medida de desempenho e a diferenciam de métrica, definindo medida como o instrumento utilizado para quantificar a eficiência e/ou eficácia da ação. A métrica é definida pelos três elementos indispensáveis e necessários: a medida, um padrão de desempenho e as consequências relativas ao estar abaixo ou acima do padrão.

Nesse sentido, para Melnyk *et al.* (2014) a métrica é mais que uma unidade de medida. Ela é o estabelecimento de um grau de referência (sinalizando um bom ou mau desempenho) das possíveis consequências de ocorrência. Behn (2003) ressalta que os gestores públicos precisam de uma medida e de um padrão de desempenho para guiá-los no alcance dos propósitos organizacionais.

Além das medidas individuais, Neely *et al.* (1995) apontam que o conjunto dessas medidas formam um sistema de avaliação de desempenho que interage com o ambiente, estruturado em níveis, conforme mostra a Figura 3.

As medidas individuais podem ser categorizadas de várias maneiras e precisam considerar a estratégia organizacional, pois influenciam o comportamento das pessoas. Nesse nível, devem-se identificar, inicialmente, quais são as medidas, para que elas são usadas, seu custo e quais os benefícios da sua utilização (Neely, 1999; Neely *et al.*, 1995).



**Figura 3.** Framework de Avaliação de Desempenho

Fonte: Traduzido de "Performance measurement system design: a literature review and research agenda" (p. 81), Neely *et al.* (1995). *International Journal of Operations & Production Management*, 15(4), 80-116.

No nível superior, o sistema de avaliação de desempenho deve ser analisado no sentido de envolver todos os elementos apropriados (internos, externos, financeiros e não financeiros), relacionar as medidas com o grau de melhoria e com os objetivos de longo e curto prazos, integrar e verificar se existem conflitos entre as medidas. No nível mais amplo, o sistema deve interagir com o ambiente interno e externo, e a análise visa a identificar se as medidas reforçam as estratégias, correspondem à cultura organizacional, são consistentes com a estrutura de reconhecimento e recompensa, concentram-se na satisfação do cliente e apontam o que a concorrência está fazendo (Neely *et al.*, 1995).

Com relação à interação dos sistemas com o ambiente, Neely *et al.* (1995) dividem os ambientes em interno, que se refere à organização, e externo, que se refere ao mercado, dividido em clientes, fornecedores e concorrentes. Internamente, o sistema de avaliação de desempenho deve ser consistente com a cultura organizacional. Externamente, os autores buscaram, por meio de pesquisas de opinião e pesquisas de mercado, identificar os interesses dos clientes, com forte ponderação à sua satisfação.

Neely *et al.* (1995) apontam que o *benchmarking* é a técnica que pode ser utilizada para avaliar o desempenho do concorrente, ainda que não tenha sido criada para isso. O *benchmarking* também pode assinalar oportunidades de melhoria e facilitar o aprendizado (Behn, 2003; Neely, 1999; Neely *et al.*, 1995).

Um sistema de avaliação de desempenho equilibrado deve proporcionar aos gerentes informações sobre todos os *stakeholders*, sejam eles internos, externos ou ambos (Melnyk *et al.*, 2014; Neely *et al.*, 1995). Para Choong (2014) as partes interessadas não são consideradas nos sistemas de avaliação de desempenho, causando problemas em relação às metas e ao tipo de informação que será

comunicada aos *stakeholders*. Identificar corretamente as partes interessadas evita esses problemas.

O estudo de Behn (2003) define que as medidas de desempenho podem ser utilizadas para atingir oito propósitos organizacionais distintos. Os propósitos são: avaliar, controlar o comportamento, orçar, motivar, promover, celebrar, aprender e melhorar, sendo melhorar o desempenho o único objetivo real. A Tabela 2 ressalta as características que as medidas de desempenho devem ter para se alcançar os objetivos organizacionais.

Tabela 2

**Características das Medidas de Desempenho Com Base no Objetivo Pretendido**

<b>Objetivos</b>	<b>Para usar as medidas de desempenho para alcançar o objetivo, o gerente público precisa</b>
Avaliar	De algum tipo de resultado desejado para comparar os dados e, assim, julgar o desempenho.
Controlar	Primeiro estabelecer o padrão comportamental ou de entrada desejado, a partir do qual medirá o desvio individual ou coletivo.
Orçar	Ter uma ideia do que é um nível de eficiência bom, aceitável ou ruim.
Motivar	Ter alguma noção do que são metas razoáveis e significativas.
Promover	Entender com o que o público se importa.
Celebrar	Discernir os tipos de realizações que os funcionários e colaboradores acham que valem a pena comemorar.
Aprender	Ser capaz de detectar desenvolvimentos inesperados (e significativos) e antecipar uma ampla variedade de comportamentos organizacionais, humanos e sociais comuns.
Melhorar	Entender (ou prever) como suas ações afetam o comportamento dentro da 'caixa preta' das pessoas que contribuem para os resultados desejados.

**Nota.** Adaptado de "Why measure performance? Different purposes require different measures" (p. 593). Fonte: Behn, R. D. (2003). *Public Administration Review*, 63(5), 586-606.

Franco-Santos *et al.* (2007) buscam, por meio de uma revisão da literatura, definir as características principais de um sistema de avaliação de desempenho. Eles consideram as seguintes bases como essenciais para definir a AD: as características do sistema de avaliação de desempenho, os papéis que desempenha e os processos que fazem parte dele.

Tendo em vista que as características são elementos ou propriedades, os papéis são os propósitos ou funções executados pelo sistema de avaliação de desempenho, e os processos são a série de ações que se combinam para constituirlo (Franco-Santos *et al.*, 2007). É possível traçar um paralelo entre os papéis e os oito propósitos de Behn (2003). Pode-se dizer que são duas as características necessárias: medidas de desempenho e infraestrutura de apoio. Claramente, as medidas são um requisito necessário, ou seja, não existe sistema de avaliação de desempenho sem elas. Já a infraestrutura de apoio pode variar de métodos manuais para sistemas informatizados (Franco-Santos *et al.*, 2007).

Nudurupati *et al.* (2011) complementam que os sistemas de informação gerenciais são fundamentais para a implementação do sistema de avaliação de desempenho. Choong (2014) ressalta o cuidado que se deve ter ao se caracterizar os sistemas de avaliação de desempenho. Segundo a Teoria Geral dos Sistemas, de Bertalanffy (1968 como citado em Choong, 2014), um sistema é caracterizado pela sinergia dos subsistemas.

Os sistemas de avaliação de desempenho podem ser divididos em três ciclos principais: desenho, implementação e uso, sendo a atualização incluída na última fase (Bourne *et al.*, 2000; Nudurupati *et al.*, 2011). Os dois requisitos da fase de projeto são: identificar os principais objetivos e desenhar as medidas de desempenho. Para Bourne *et al.* (2000, p. 758) na fase de aplicação “os sistemas e procedimentos são postos em prática para coletar e processar os dados” e apresentá-los de forma significativa. A utilização das medidas é dividida em avaliar a implementação da estratégia e desafiar os pressupostos estratégicos. Para completar esse ciclo de vida, deve-se ainda atualizar o sistema (Bourne *et al.*, 2000).

Bititci *et al.* (2015) apontam que existem dois subsistemas: mensuração, voltada à preocupação com aspectos técnicos e estruturais, como, por exemplo, qual é a medida, como será mensurada, quais serão o alvo e as metas; e gerenciamento, no qual a preocupação se volta aos aspectos culturais e comportamentais, a como a informação será recebida pelas pessoas. Melnyk *et al.* (2014) complementam que a mensuração está na parte operacional. Os ciclos de vida, estudados por Bourne *et al.* (2000), são a fase de construção e uso do modelo, enquanto o gerenciamento está conectado à retroalimentação do sistema, completando as fases dos ciclos de vida.

A retroalimentação do sistema está conectada à atualização, que deve ocorrer sempre em três situações: alterações no ambiente, estratégia e cultura organizacional. Assim, as pessoas que operam o sistema devem estar prontas para procederem o ajuste das medidas, de forma a evitar os conflitos internos (Melnyk *et al.*, 2014).

O processo de atualização é dividido em quatro processos: rever as metas, desenvolver medidas, rever as medidas e, por fim, alterar a estratégia. Portanto, é no processo de atualização que se pode desafiar os pressupostos estratégicos. Nesse processo, é posto à prova se as medidas estão realmente alinhadas aos objetivos estratégicos, sendo possível identificar esse alinhamento com a criação de um mapa de causa e efeito (Bourne *et al.*, 2000; Matos *et al.*, 2019).

Ghalayini e Noble (1996) afirmam que os sistemas de avaliação de desempenho, desenvolvidos até então, apresentavam limitações, pois eram constituídos como ferramentas de melhoria e controle, sem considerar a integração com a melhoria contínua. Não era possível especificar qual objetivo devia ser cumprido. Eles não eram dinâmicos, pois não permitiam a revisão sistemática. Não havia preocupação com o futuro e tampouco ferramentas que enfatizavam a otimização global. Tais sistemas não ressaltavam a importância do tempo, e não havia ferramentas para o chão de fábrica.

Para Bourne *et al.* (2000) os três obstáculos principais à implantação de um sistema de avaliação de desempenho são: resistência às medidas de desempenho, que ocorre nas fases de desenho e utilização; relacionado aos sistemas de informação na fase de execução das medidas; e distração da alta gerência entre as

fases de desenho e execução. Já Melnyk *et al.* (2014) apontam a falta de transparência no processo de tomada de decisão, advinda da utilização do SAD, como um obstáculo que prejudica a aprendizagem do processo.

Camp e Braet (2016) aprofundam o tema, e identificam vários fatores que podem levar ao fracasso do SAD. No nível das métricas, eles encontraram 13 falhas associadas às medições, parâmetros e indicadores de desempenho. No segundo nível (*framework*), estão 9 falhas referentes à estrutura e ao modelo, basicamente ao processamento das medidas. Por fim, no terceiro nível, referente às decisões dos gestores, estão 14 falhas atribuídas à implementação do SAD, resultando em 36 advertências.

As falhas no nível das métricas, mostradas na Tabela 3, estão relacionadas a perguntas sobre coleta de dados: O quê? Como? Quanto? Quando medir? Nesse nível, falhas surgem da seleção de métricas em razão da praticidade; de como utilizar métricas de outras empresas; de selecionar métricas rápidas, fáceis e econômicas de medir; de ter dificuldades de mensurar intangíveis, incertezas ou devido à restrição de recursos simples. As métricas ativas devem ser selecionadas com cuidado e validadas continuamente (Camp & Braet, 2016).

A estrutura refere-se ao conjunto de técnicas, modelos e métodos utilizados para analisar medidas. A primeira falha é referente ao escopo impreciso, que leva a estruturas e modelos ineficientes. Devido à heterogeneidade das diferentes dimensões, surge uma quantidade elevada de modelos disponíveis que nem sempre conversam entre si e que levam a estimativas diferentes de desempenho. O conhecimento pode ajudar no combate às falhas apresentadas.

Por fim, segundo Camp e Braet (2016), a principal falha, no nível gerencial, é referente à falta de apoio dos gestores. Disso podem ser desencadeados outros problemas, como falta de recursos financeiros, humanos e tecnológicos para a implementação do SAD, de envolvimento no acompanhamento e de ações reativas/proativas.

A gerência pode colaborar na implementação bem-sucedida do SAD, ao combater essas falhas e preparar a cultura interna para o seu uso diário. Outras faltas são apontadas, tais como desalinhamento estratégico, deixar de atualizar o sistema com frequência e sistema de recompensas inadequado, pode se referir às recompensas em relação ao desempenho superior em um indicador (Camp & Braet, 2016).

Neely (1999) afirma que o campo de pesquisa da avaliação de desempenho é multidisciplinar, e muitos ganhos podem surgir se áreas de pesquisa diferentes trabalharem em conjunto. Bititci *et al.* (2012) salientam que o principal desafio é desenvolver uma compreensão integrada e holística da avaliação de desempenho como um fenômeno social, que permita a aprendizagem em redes autopoiéticas. Por lidar com pessoas, a gestão da mudança pode desempenhar papel fundamental na implantação bem-sucedida dos ciclos de vida dos sistemas de avaliação de desempenho (Nudurupati *et al.*, 2011).

Tabela 3

**Taxonomia das Falhas no Nível Métrico**

Cód.	Falha nas métricas	Descrição
M1	Falta de definição clara, única e transparente	Devido à grande diversidade de papéis do próprio SAD e à necessidade de incorporar múltiplas dimensões, é essencial ter definições operacionais bem documentadas, não ambíguas. Os autores pedem para declarar claramente o item (o quê?), a unidade (como?) e o valor (por quê?).
M2	Transposta de outras empresas	Devido à falta de conhecimento, motivação e/ou recursos financeiros, as empresas demonstram um comportamento de “copiar e colar” de empresas (geralmente) maiores. No entanto, métricas e sistemas <i>off-shelf</i> representam flexibilidade insuficiente.
M3	Selecionada em acessibilidade e disponibilidade	O custo total verdadeiro do SAD não deve ser subestimado. Os principais contribuintes são a construção e manutenção do sistema, coleta e estruturação de informações e custo de tempo de todos os envolvidos. Portanto, as empresas recorrem ao que é rápido, fácil e econômico de medir. Consequentemente, os vieses estão presentes, prejudicando o sistema e levando a decisões incorretas.
M4	Quantidade desequilibrada	Oblíquidade ou simplificação excessiva é obtida se muitos ou poucos indicadores forem monitorados, respectivamente. O primeiro é promovido pela incapacidade de descartar métricas (falta de métricas dinâmicas) ou falta de foco. O último pode ser resultado do fracasso M3 ou de foco limitado nas métricas da produção financeira.
M5	Foco dominante em métricas financeiras	Historicamente, o foco dominante é em métricas financeiras. No entanto, os autores sugerem que é mais interessante medir o que é necessário (entrada e processo) para lucrar. As métricas da produção financeira complicam a análise por serem indicadores atrasados, insinuando o que foi definido como perda. Além disso, é difícil alocar a produção financeira à entrada, negando possíveis sinergias.
M6	Relação desequilibrada entre métricas qualitativas e quantitativas	Dificuldades na obtenção de números ou falta de cálculos avançados fazem com que se recorra ao uso excessivo de métricas qualitativas, que representem uma abordagem mais subjetiva, levando a resultados menos robustos. O excesso de ênfase nas métricas quantitativas, erroneamente, se tornou sinônimo de “supermonetização”, ou seja, tentativa de medir tudo com métricas financeiras (M5).
M7	Dificuldades de mensuração de intangíveis	A maioria das empresas luta para medir os atributos não físicos de forma robusta. Sua abordagem é consultar usuários (especialistas), medir despesas e recursos de P&D e ampliar as estatísticas de patentes. Uma interação interessante está sendo criada entre intangíveis, aproximações qualitativas e quantitativas.
M8	Conjunto incompleto	As múltiplas dimensões do SAD agravam o processo de seleção de métricas, pois se busca um grau de cobertura suficiente: métricas internas e externas, métricas de entrada, processo e saída. Autores diversos propõem uma estrutura diferente, como medidas do <i>balanced scorecard</i> sob quatro perspectivas: finanças, processos internos de negócios, aprendizagem e crescimento e clientes.
M9	Perigo de métricas se tornarem alvos	“As medidas feitas para provar raramente melhorarão”. A fixação da medição obstruirá a visão clara do objetivo e poderá estimular resultados não desejáveis. Essa falha se intensifica quando associada a um sistema de recompensa inadequado.

Tabela 4

**Taxonomia das Falhas no Nível Métrico**

(conclusão)

M10	Falta de métricas robustas	O valor da métrica deve ser baseado em valores mensuráveis, e não é fácil de ser manipulado ou manipulado pela “equipe de provedores”. Combinar robustez, especificidade e sensibilidade é uma tarefa difícil.
M11	Falta de métricas objetivas	Recorre-se a métricas subjetivas devido a várias limitações: falta de dados disponíveis (M3), dificuldades de tradução de bens intangíveis (M7), incerteza (M12) e recursos limitados (tempo, pessoas e dinheiro). Diferentes autores alertam para deficiências de modelos arbitrários de pontuação.
M12	Incerteza no início de um projeto	A incerteza inicial impede a análise das métricas encontradas mais adiante no processo de produção, como, por exemplo, durante o estágio de geração ou seleção de ideias, pouco se sabe sobre os custos ou vendas reais de produção e distribuição.
M13	Uso indevido de métricas deterministas	“É melhor estar mais ou menos certo do que precisamente errado”. Frequentemente, precisão e valores determinísticos são tratados como sinônimos. Deve-se cuidar da “falha das médias”, descartando-se muitas informações e apenas mantendo as estimativas pontuais, pois elas propõem um senso de precisão confortável, mas falso.

**Nota.** Adaptado de "Taxonomizing performance measurement systems' failures" (pp. 676-679). Fonte: Camp, J. van, & Braet, J. (2016). *International Journal of Productivity and Performance Management*, 65(5), 676-679.

Lebas (1995) conclui que um sistema de gestão de desempenho robusto é aquele que utiliza medidas que dão autonomia aos indivíduos dentro da amplitude de controle, refletem a causa e o efeito dos relacionamentos, capacitam e envolvem os indivíduos, criam uma base para discussão e apoiam a melhoria contínua, além da tomada de decisão.

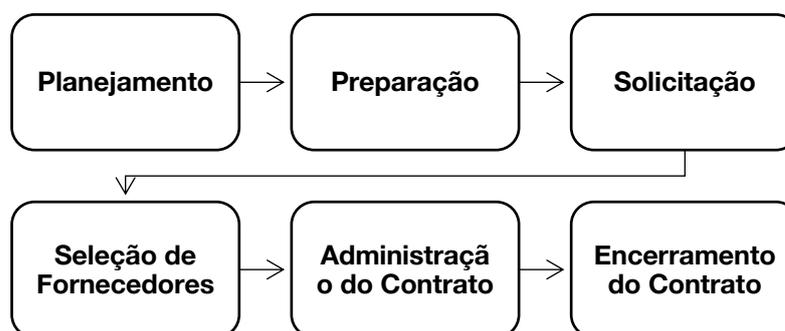
### Avaliação de desempenho de compras públicas

Os primeiros estudos de avaliação de desempenho de compras datam de 1928, quando foi publicado o estudo feito por Gushée e Boffey, com foco no setor privado e enfatizando a necessidade de medir a eficiência e a eficácia da aquisição (Patrucco *et al.*, 2016; Schiele, 2005). Apesar das diferenças entre eles, verifica-se que o setor público busca acompanhar o sucesso do setor privado ao adotar métodos já consagrados, como as compras eletrônicas ou o *balanced scorecard* (Brignall & Modell, 2000; Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Rendon, 2008). Em geral, os contratos públicos atendem a uma regulamentação maior, devendo cumprir mais regras que são impostas pela legislação (Srivastava & Agrahari, 2017).

Schapper, Malta e Gilbert (2006) afirmam que reformas administrativas cíclicas e sucessivas impactam o ambiente de compras públicas, ora buscando melhor desempenho em direção à responsabilização pelos resultados, ora buscando mais conformidade à regulamentação devido às falhas que ocorreram no processo anterior. Rendon (2008) afirma que o incentivo legislativo estimulou as organizações públicas a buscarem melhoria no desempenho. Percebe-se, portanto, que a atividade

de compras tem forte influência dessa regulamentação (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017).

Além da forte influência da regulamentação nas compras públicas no âmbito local, entende-se compras públicas como um processo para aquisição de bens, obras e serviços para os departamentos governamentais (Srivastava & Agrahari, 2017). Thai (2008 como citado em Srivastava & Agrahari, 2017) aponta a identificação da necessidade, a seleção dos fornecedores, a adjudicação e gestão dos contratos como fases do processo. Rendon (2008) elenca planejamento, preparação, solicitação, seleção dos fornecedores, administração do contrato e encerramento do contrato, conforme revela a Figura 4.



**Figura 4.** Fases do Processo de Compras Públicas

O processo de compras públicas é dinâmico e envolve licitantes, reguladores e agências de aquisições (Srivastava & Agrahari, 2017). Além disso, Schiele (2005) assevera que o setor de compras possui uma função estratégica devido a esse contato com *stakeholders*, internos e externos.

Os *stakeholders* internos definidos na literatura são: gestores (Chomchaiya & Esichaikul, 2016; Rendon, 2008; Schiele, 2005; Srivastava & Agrahari, 2017); usuários do serviço, também denominados clientes internos (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Rotchanakitumnui, 2013; Schiele, 2005); equipe de suporte ou encarregado de compras (Chomchaiya & Esichaikul, 2016; Rotchanakitumnui, 2013; Srivastava & Agrahari, 2017); e auditores e diretores financeiros (Chomchaiya & Esichaikul, 2016).

Já os *stakeholders* externos são: empresas licitantes ou fornecedores (Gardenal, 2013; Patrucco *et al.*, 2016; Rotchanakitumnui, 2013; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017); auditores externos (Chomchaiya & Esichaikul, 2016; Schapper *et al.*, 2006); e agências multilaterais de desenvolvimento e órgãos reguladores (Srivastava & Agrahari, 2017). Além da relação do Departamento de Compras com as partes interessadas, verifica-se o impacto das compras públicas na sociedade em geral em virtude do alto volume de recursos despendidos (Gardenal, 2013; Patrucco *et al.*, 2016; Rendon, 2008; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017). Por isso, entende-se que as aquisições desempenham papel central nas organizações públicas, surgindo a necessidade de medir e melhorar os processos de compras, promovendo, assim, a realização dos objetivos públicos

gerais (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016; Rendon, 2008; Schiele, 2005; Srivastava & Agrahari, 2017).

Resumidamente, pode-se dizer que os objetivos das compras públicas estão ligados ao atendimento dos padrões da comunidade (Patrucco *et al.*, 2016; Schapper *et al.*, 2006), ao uso das verbas públicas de forma eficaz e eficiente (Gardenal, 2013; Patrucco *et al.*, 2016; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017), à conformidade legislativa (Patrucco *et al.*, 2016; Rendon, 2008; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017) e ao cumprimento das metas das políticas do governo e/ou organizacionais (Patrucco *et al.*, 2016; Rendon, 2008; Rotchanakitumnuai, 2013; Schapper *et al.*, 2006; Schiele, 2005).

Relaciona-se ao primeiro objetivo a promoção da concorrência por meio do tratamento justo e equitativo dos licitantes (Gardenal, 2013; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017), a responsabilização (Chomchaiya & Esichaikul, 2016; Gardenal, 2013; Rendon, 2008; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017) e a transparência (Chomchaiya & Esichaikul, 2016; Gardenal, 2013; Patrucco *et al.*, 2016; Rotchanakitumnuai, 2013; Schapper *et al.*, 2006; Srivastava & Agrahari, 2017). Já a redução dos custos está relacionada ao uso dos recursos públicos (Chomchaiya & Esichaikul, 2016; Gardenal, 2013; Patrucco *et al.*, 2016; Rotchanakitumnuai, 2013; Schapper *et al.*, 2006; Schiele, 2005; Srivastava & Agrahari, 2017).

Alcançar esses objetivos depende da racionalização dos gastos públicos, que será proporcionada por gestão e controle da função de compras, podendo, assim, contribuir para o desempenho organizacional, tornando os sistemas de avaliação de desempenho de compras públicas relevantes (Gardenal, 2013; Patrucco *et al.*, 2016; Srivastava & Agrahari, 2017). Apesar da sua importância, Patrucco *et al.* (2016) apontam que o tema tem sido pouco investigado.

## Procedimentos metodológicos

### Enquadramento metodológico

Com o intuito de atender aos objetivos formulados, esta pesquisa caracteriza-se por ser exploratória e descritiva. A pesquisa exploratória tem como objetivo “conhecer as características de um fenômeno para procurar, posteriormente, explicações das causas e consequências de dito fenômeno” (Richardson, 1999, p. 326). Isso ocorre neste trabalho por meio da seleção, reflexão e interpretação do significado das variáveis analisadas do portfólio bibliográfico (PB). Já a pesquisa descritiva pretende “descrever sistematicamente um fenômeno ou área de interesse” (Richardson, 1999, p. 326), de forma detalhada e objetiva, sendo usada quando são identificados os ‘destaques’ das características investigadas do PB.

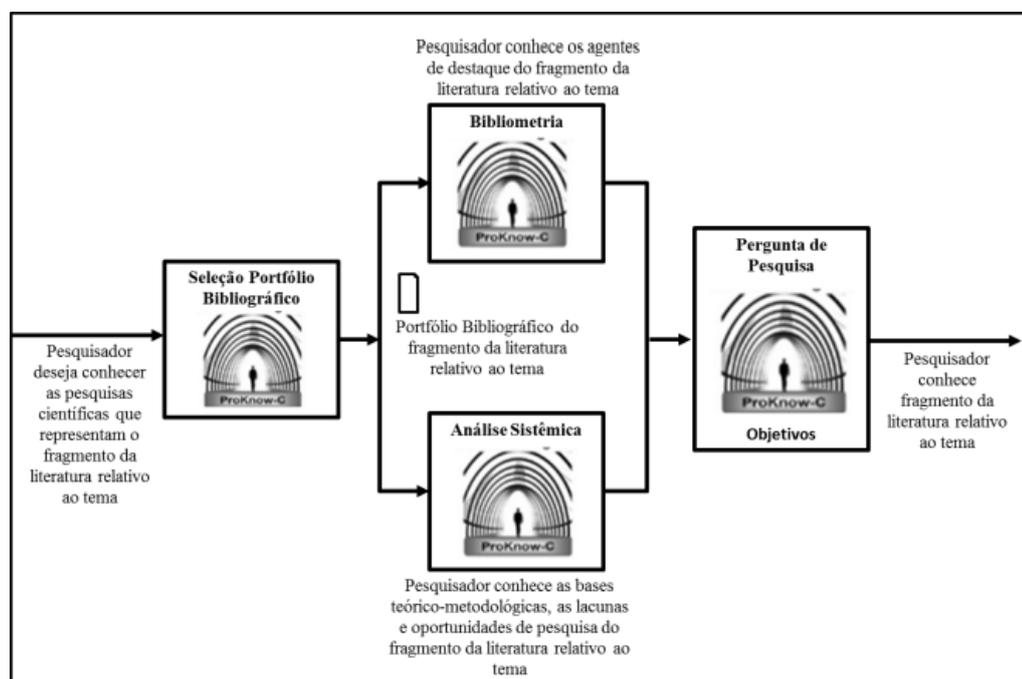
Em relação aos procedimentos técnicos, foram utilizadas a pesquisa bibliográfica, pois envolveu a análise dos dados extraídos dos artigos científicos, e a pesquisa-ação, pelo envolvimento dos pesquisadores na geração do conhecimento para dar significado aos ‘destaques’ das características investigadas e evidenciá-los.

Quanto à coleta de dados, foram coletados dados primários e secundários. Os dados primários foram utilizados em virtude das delimitações postas pelos pesquisadores para a seleção do PB, e os dados secundários por serem coletados nos artigos científicos que compõem o PB.

Por fim, a abordagem do problema é qualitativa. Segundo Richardson (1999, 79), o uso da abordagem qualitativa se justifica “por ser uma forma adequada para entender a natureza de um fenômeno social”. Isso foi exatamente o que os pesquisadores propuseram, uma vez que eles pretendiam conhecer as características do fragmento da literatura acerca das compras públicas sob a perspectiva da avaliação de desempenho, e interpretá-las dando significado aos ‘destaques’ identificados.

### Instrumento de intervenção

O instrumento de intervenção *Knowledge Development Process-Constructivist (ProKnow-C)* tem o objetivo de conduzir o pesquisador no caminho da geração de conhecimento sobre o fragmento da literatura científica de seu interesse por meio de um processo estruturado e sistematizado. O processo é realizado por meio de quatro etapas: seleção do portfólio bibliográfico que representa um fragmento da literatura, análise bibliométrica do PB, análise sistêmica do PB e identificação de lacunas e oportunidade que conduzem às propostas de perguntas de pesquisas futuras (Ensslin *et al.*, 2015; Thiel *et al.*, 2017; Valmorbida *et al.*, 2016;). Tal processo é mostrado na Figura 5.



**Figura 5.** Etapas do *ProKnow-C*

Fonte: Extraído de "Rankings universitários mundiais: Que dizem os estudos internacionais?" (p. 12), Valmorbida, S. M. I., Ensslin, S. P. R., Ensslin, L., & Ripoll-Feliu, V. M. (2016). *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 14(2), 5-29.

## Procedimento de coletas de dados

Com uso do instrumento de intervenção *ProKnow-C*, iniciou-se a seleção do portfólio bibliográfico. Esse PB é composto por um fragmento da literatura referente ao tema, julgado pelos autores como representativo da literatura consultada para realizar as análises pertinentes. Assim, a partir do instrumento de intervenção, dá-se início a busca de artigos científicos em bases de dados com suficiente representatividade, como indicado por Matos *et al.* (2019), filtrando os resultados segundo critérios de relevância e reconhecimento científico para composição do PB.

A consulta às bases de dados disponíveis no Portal de Periódicos da Capes, com base nos eixos de pesquisa selecionados, foi feita entre os dias 29 e 31 de março de 2018. Foram delimitados apenas artigos publicados em *journals*, em língua inglesa, a partir do ano 2000. Os procedimentos para a seleção do PB, bem como as bases consultadas e as palavras-chave utilizadas, estão ilustrados na Figura 6.

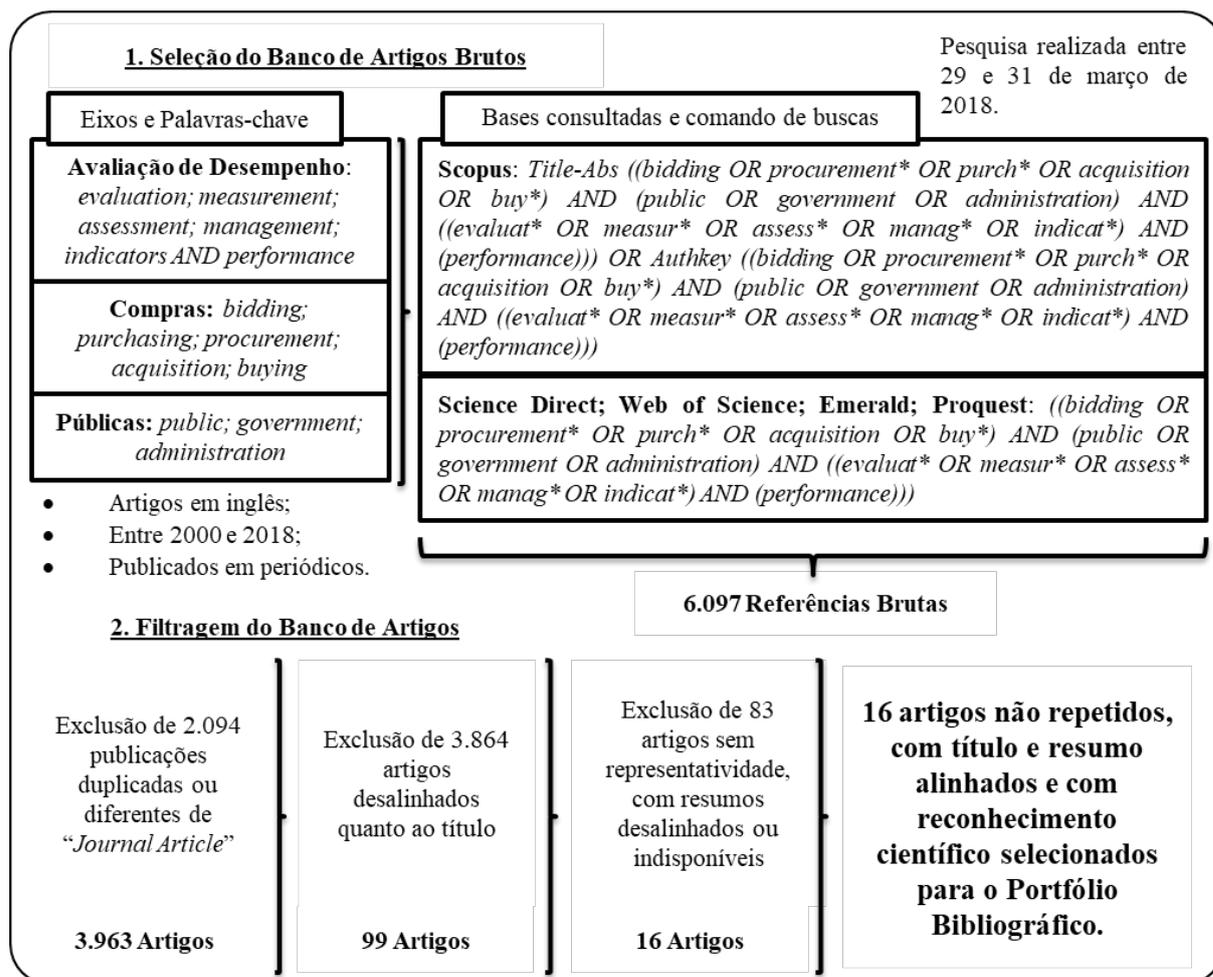
Foram localizados nas bases de dados 6.097 referências, que foram importadas para a ferramenta Endnote X7. Com o auxílio da ferramenta foram excluídas as referências duplicadas, que resultaram em 3.963 artigos. Examinou-se, então, o título desses artigos quanto ao alinhamento ao tema da presente pesquisa e foram excluídos os títulos claramente desalinhados, que, na percepção dos pesquisadores, não contribuiriam para o seu desenvolvimento. Referências duplicadas que não foram automaticamente excluídas pela ferramenta também foram eliminadas.

Para a seleção dos artigos mais representativos, determinou-se no Google Scholar o número de citações dos 99 artigos do banco de artigos com títulos alinhados. Em seguida, os artigos foram ordenados pelo número de citações e participação em relação ao total de citações. Fixou-se a representatividade desejada em 94,8% do total de citações, em que foram selecionados os 50 artigos com mais de 11 citações. Os 49 artigos restante foram separados no Repositório P.

A partir da leitura do resumo dos 50 artigos, foram identificados 22 artigos com título e resumo alinhados ao tema de avaliação de desempenho de compras públicas e com reconhecimento científico, que formaram o Repositório A. Foram identificados os 55 autores desses artigos.

Nos artigos do Repositório P foram identificados 31 artigos recentes, publicados nos três anos anteriores a pesquisa. Buscou-se nos artigos não recentes do Repositório P se havia algum artigo escrito pelos autores do Repositório A, e foi identificado um artigo. A partir da leitura do resumo dos 32 artigos selecionados, foram identificados 10 artigos com título e resumo alinhados ao tema de pesquisa e com reconhecimento científico em potencial, formando o Repositório B.

Após a leitura integral dos 32 artigos (Repositórios A e B) e com o desenvolvimento do processo mostrado na Figura 6, foram selecionados 16 artigos relacionados ao fragmento da literatura que trata da avaliação de desempenho de compras públicas – os artigos selecionados estão listados na Tabela 4. Buscou-se ainda, nas referências, novos artigos que pudessem ser incorporados ao PB. Porém, nenhum foi selecionado em virtude de os artigos das referências não se enquadrarem nas delimitações anteriormente utilizadas na seleção do PB original (alinhamento ao tema, palavras-chave e período).



**Figura 6.** Resumo do Processo de Seleção do Portfólio Bibliográfico

### Procedimento para análise dos dados

Com base no portfólio bibliográfico, foi feita a análise bibliométrica, dividida em análise das características básicas e características das avançadas, como recomendado por Thiel *et al.* (2017). As características básicas delimitadas para investigação desse PB são: os autores prolíficos no tema e as redes de colaboração (coautores); o foco que tem merecido atenção dos pesquisadores; e a finalidade para a qual a avaliação de desempenho tem sido utilizada nos estudos.

As características avançadas delimitadas para investigação desse PB são: os níveis em que o sistema de avaliação de desempenho está sendo abordado (Neely *et al.*, 1995); os propósitos organizacionais para utilizar as medidas de desempenho (Behn, 2003); e as falhas encontradas nas medidas de desempenho (Camp & Braet, 2016). Assim, com base nas características definidas para investigação, cumpre observar que somente os artigos empíricos foram analisados.

Com base no trabalho de Neely *et al.* (1995), foi observado o nível em que o sistema de AD está sendo abordado no artigo, seja em termos de medidas de desempenho individuais, de um conjunto de medidas de desempenho, da existência de interação entre o sistema de avaliação de desempenho, seja o ambiente onde a

instituição, por meio do SAD, está inserida. Tais medidas foram apresentadas na Figura 3.

No nível das medidas, foram identificadas quais as medidas de desempenho utilizadas. Complementando o conhecimento, verificou-se se existe alinhamento das medidas com a estratégia e/ou com os objetivos da instituição. No nível do sistema de avaliação de desempenho, foi investigada a sua finalidade (mensurar, diagnosticar, comparar ou informar gestores). Finalmente, no nível da interação do sistema com o ambiente, foi verificado se ela ocorre apenas no contexto interno, no externo ou em ambos.

Quanto à segunda variável avançada, Behn (2003) aponta que os gestores públicos podem usar medidas de desempenho para avaliar, controlar, orçar, motivar, promover, celebrar, aprender e melhorar, a fim de atingir o propósito organizacional, conforme detalhado na Tabela 2. Assim, pretendeu-se identificar os propósitos mais recorrentes do uso das medidas de desempenho nos estudos. A terceira variável avançada investigada está relacionada com a identificação das falhas das medidas de desempenho utilizadas nos artigos, com base na taxonomia criada por Camp e Braet (2016), mostrada na Tabela 3.

Por fim, elaborou-se a pauta de pesquisa com as lacunas identificadas, utilizando-se os resultados apurados na análise bibliométrica, composta por proposição de contribuições para o meio acadêmico e para o meio gerencial.

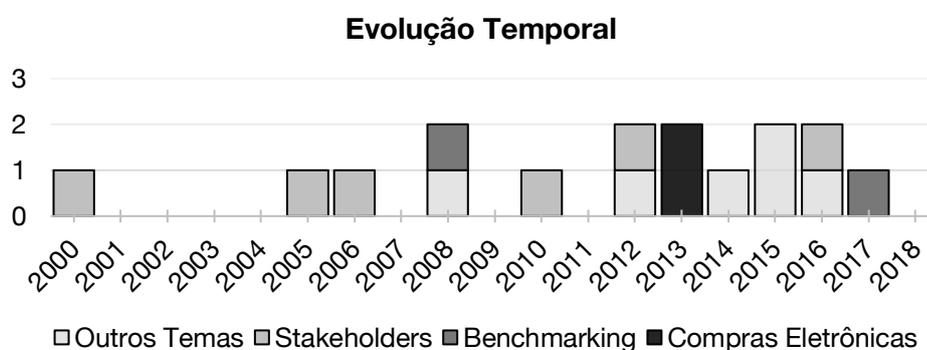
## Análise e discussões

### Análise bibliométrica das características básicas

Com base na análise bibliométrica das variáveis básicas do PB, buscou-se identificar os pesquisadores com trajetória de pesquisa na literatura sobre avaliação de desempenho de compras públicas. Considerando que no PB não há qualquer autor com mais de uma publicação, fez-se uma análise das demais publicações dos autores. Entretanto, não foi localizado qualquer autor prolífico nessa área. Tal fato é corroborado nos artigos do fragmento de literatura selecionado, nos quais é mencionada a falta de interesse acadêmico em pesquisar compras públicas sob a perspectiva da avaliação de desempenho (Chomchaiya & Esichaikul, 2016; Patrucco *et al.*, 2016). Percebe-se, pela análise do histórico de publicação dos autores do PB, que pesquisadores de diferentes áreas somaram conhecimento a esse fragmento ao investigarem e disseminarem os resultados encontrados (Galera *et al.*, 2008; Goh, Elliott & Richards, 2015; Kim, Olivares & Weintraub, 2014).

Em virtude do número de publicações dos autores no PB, não foi possível estabelecer a existência de redes de colaboração. Entretanto, percebeu-se que em dois artigos de autores diferentes, quais sejam, os de Gardenal (2013) e Patrucco *et al.* (2016), a avaliação de desempenho de compras públicas foi investigada na região da Lombardia, Itália, provavelmente pela utilização de ferramentas específicas para se avaliar o desempenho pela Agência Regional de Compras (*Regional Procurement Agency*) desde 2008.

Na sequência, considerou-se relevante conhecer o foco pelo qual as compras públicas, sob a perspectiva da avaliação de desempenho, têm merecido atenção dos pesquisadores. A Figura 7 sumariza o resultado.



**Figura 7.** Foco dos Artigos ao Longo dos Anos

Fonte: dados da pesquisa.

Chama atenção o fato de seis artigos do PB tratarem do tema de compras públicas sob a perspectiva da avaliação de desempenho com foco nos *stakeholders*, tanto internos quanto externos. Além dos *stakeholders*, é possível observar, na Figura 6, que dois artigos tratam do tema com enfoque no *benchmarking* e dois tratam de compras eletrônicas. Os demais foram agrupados e tratam da avaliação de desempenho em um local específico, maturidade, avaliação de leilões combinatórios e contratação de obras.

Apesar da importância da informação advinda do *benchmarking* para a avaliação de desempenho na análise do ambiente externo (Neely *et al.*, 1995), cumpre destacar que não foi observado o uso do *benchmarking* nas pesquisas apresentadas nos dois artigos. Nelas os autores desenvolveram ou utilizaram indicadores genéricos para tornar possível a comparação entre os desempenhos identificados (Galera *et al.*, 2008; Srivastava & Agrahari, 2017).

Após investigar o foco, verificou-se a finalidade do uso da avaliação de desempenho nos artigos. A Tabela 4 apresenta o resultado encontrado e a informação de natureza teórica ou empírica dos artigos do PB.

Cabe destacar que nos quatro artigos teóricos não foi desenvolvido um sistema de avaliação de desempenho de um caso específico, motivo pelo qual não foi possível investigar a finalidade. No caso dos artigos empíricos, a análise da finalidade da avaliação de desempenho não se aplica, pois Galera *et al.* (2008) buscaram identificar no seu artigo os obstáculos para a implantação de indicadores de desempenho, e o artigo de Goh *et al.* (2015) segue em linha semelhante em relação aos desafios de implementação.

Tabela 5  
Finalidade da Avaliação de Desempenho

Autor (ano)	Revista	T ou E	Finalidade da Avaliação de Desempenho
Brignall e Modell (2000)	<i>Management Accounting Research</i>	Teórico	Não se aplica.
Schiele (2005)	<i>Journal of Public Procurement</i>	Empírico	Avaliar as contribuições do Departamento de Compras em processos de contratação de consultorias.
Schapper, Malta e Gilbert (2006)	<i>Journal of Public Procurement</i>	Teórico	Não se aplica.
Galera, Rodríguez e Hernández (2008)	<i>Public Management Review</i>	Empírico	Não se aplica.
Rendon (2008)	<i>Journal of Public Procurement</i>	Empírico	Avaliar o nível de maturidade do processo de aquisição.
Tillema, Mimba e Helden, (2010)	<i>Public Administration and Development</i>	Teórico	Não se aplica.
Ren, Kwaw e Yang (2012)	<i>Built Environment Project and Asset Management</i>	Empírico	Avaliar o desempenho de um tipo específico de aquisição de serviços de obras/reformas.
Reuter, Goebel e Foerstl (2012)	<i>Journal of Purchasing and Supply Management</i>	Empírico	Avaliar o impacto dos <i>stakeholders</i> nas decisões de seleção dos fornecedores (sustentabilidade versus custos).
Gardenal (2013)	<i>Journal of Public Procurement</i>	Empírico	Avaliar o desempenho das compras eletrônicas.
Rotchanakitumnuai (2013)	<i>International Journal of Physical Distribution &amp; Logistics Management</i>	Empírico	Avaliar os fatores de sucesso associados às compras eletrônicas.
Kim, Olivares e Weintraub (2014)	<i>Management Science</i>	Empírico	Avaliar o desempenho das sinergias de custos em leilões combinatórios de aquisições.
Goh, Elliott e Richards (2015)	<i>International Journal of Productivity and Performance Management</i>	Empírico	Não se aplica.
Mardale (2015)	<i>Land Forces Academy Review</i>	Teórico	Não se aplica.
Patrucco, Luzzini e Ronchi (2016)	<i>Local Government Studies</i>	Empírico	Avaliar o SAD de compras públicas de oito municípios.
Chomchaiya e Esichaikul (2016)	<i>Information Technology and People</i>	Empírico	Avaliar o desempenho de compras eletrônicas com base nos <i>stakeholders</i> internos.
Srivastava e Agrahari (2017)	<i>Economic and Political Weekly</i>	Empírico	Desenvolver indicadores padrões para comparar o desempenho do processo licitatório.

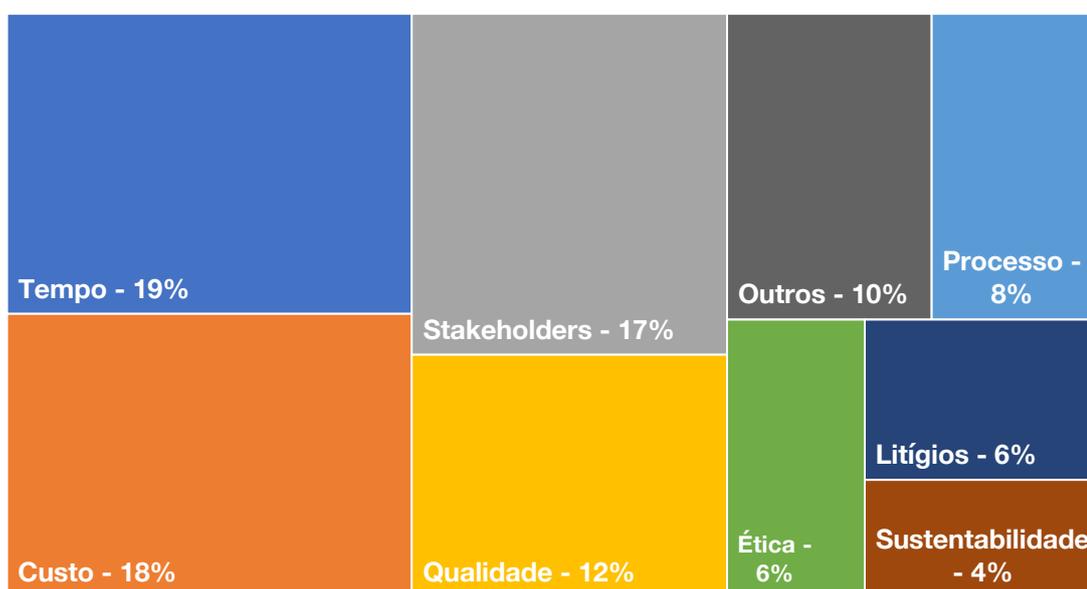
**Nota.** Fonte: dados da pesquisa.

Como era de se esperar, a finalidade da AD nos estudos foi avaliar as compras em si. A seguir, apresentam-se os achados da análise bibliométrica das

características avançadas, em que os estudos empíricos foram analisados sob o aporte teórico de autores da área de avaliação de desempenho.

### Análise bibliométrica das características avançadas

Quanto ao nível em que os sistemas de avaliação de desempenho são abordados, verificou-se que, dos 12 artigos empíricos, sete apresentaram, em conjunto, 184 medidas individuais relacionadas à AD de compras públicas (nível medidas de desempenho individuais). Essas medidas de desempenho foram agrupadas segundo as áreas de preocupações mais recorrentes, conforme mostra a Figura 8.



**Figura 8.** Áreas de Preocupações das Medidas de Desempenho

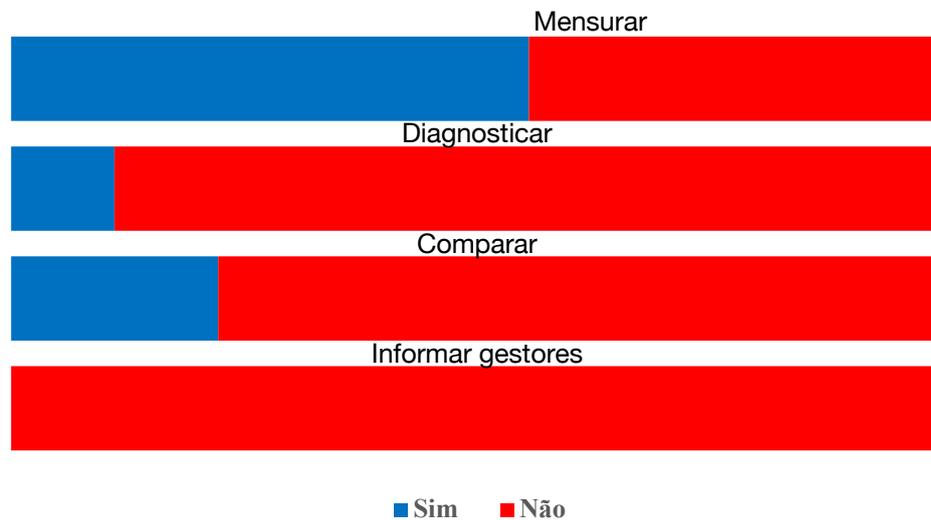
Fonte: dados da pesquisa.

Relacionadas ao custo, encontram-se as seguintes medidas individuais: economia ou desconto obtido (razão entre o valor homologado e o valor estimado), custo de aquisição do produto (valor total homologado) e redução no custo do processo (valor economizado). Vinculadas ao tempo, estão as medidas que tratam do tempo decorrido por procedimento (dias úteis), produtos entregues no prazo (percentual) e tempo gasto na análise das propostas (dias úteis). Relacionadas aos *stakeholders*, estão as medidas ligadas à concorrência e à competitividade, como o número de fornecedores que participaram do procedimento. Quanto às medidas que tratam da qualidade, há a satisfação com a qualidade do produto e o percentual de produtos recebidos com dano.

Apesar da importância das medidas estarem alinhadas à estratégia organizacional, conforme os teóricos de AD (Bititci *et al.*, 2012, 2015; Bourne *et al.*, 2000), observou-se que apenas a estrutura construída por Patrucco *et al.* (2016) considera a estratégia organizacional, ao questionar se as medidas encontradas nos

municípios estudados estavam alinhadas às metas da organização (regulatórias, comerciais ou socioeconômicas). Ou seja, os modelos construídos, constantes no PB, não são capazes de oferecer subsídios para a gestão de compras públicas.

Quanto à finalidade dos SAD, constatou-se que a mensuração é recorrente nos estudos empíricos que fazem uso de modelos, conforme Figura 9.

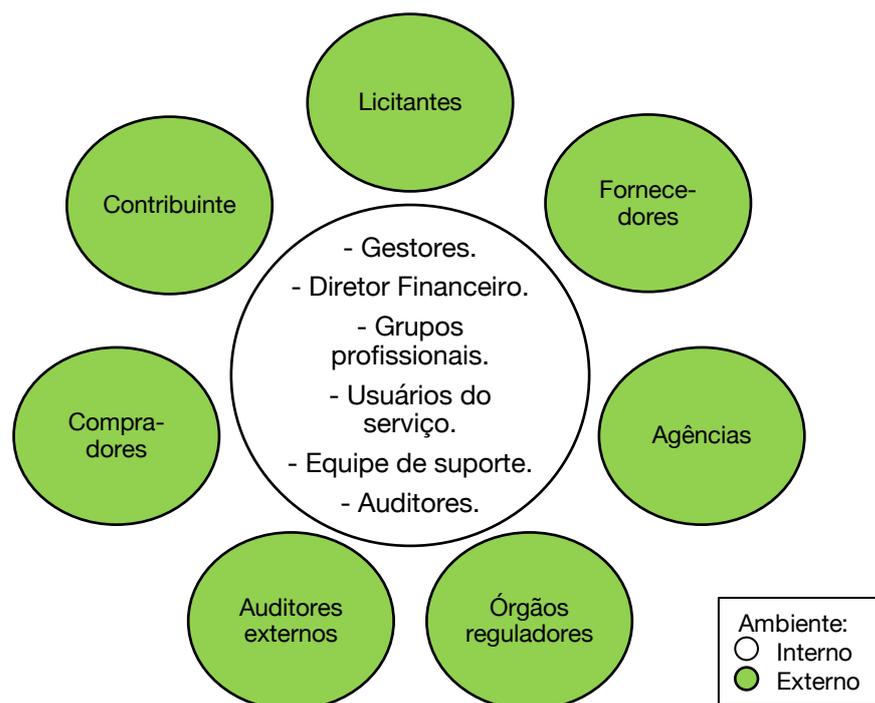


**Figura 9.** Finalidades dos Sistemas de Avaliação de Desempenho

Fonte: dados da pesquisa.

Além da falta de informação, pode-se observar que oito ferramentas não apresentam um diagnóstico do desempenho, e apenas cinco artigos mensuram o desempenho. Os dois artigos que tratam do *benchmarking* são os únicos que comparam o desempenho.

Quanto à interação do sistema com o ambiente, percebe-se que, no PB, houve preocupação em considerar a visão das partes interessadas, tanto no contexto interno quanto no externo, como se pode verificar na Figura 10.

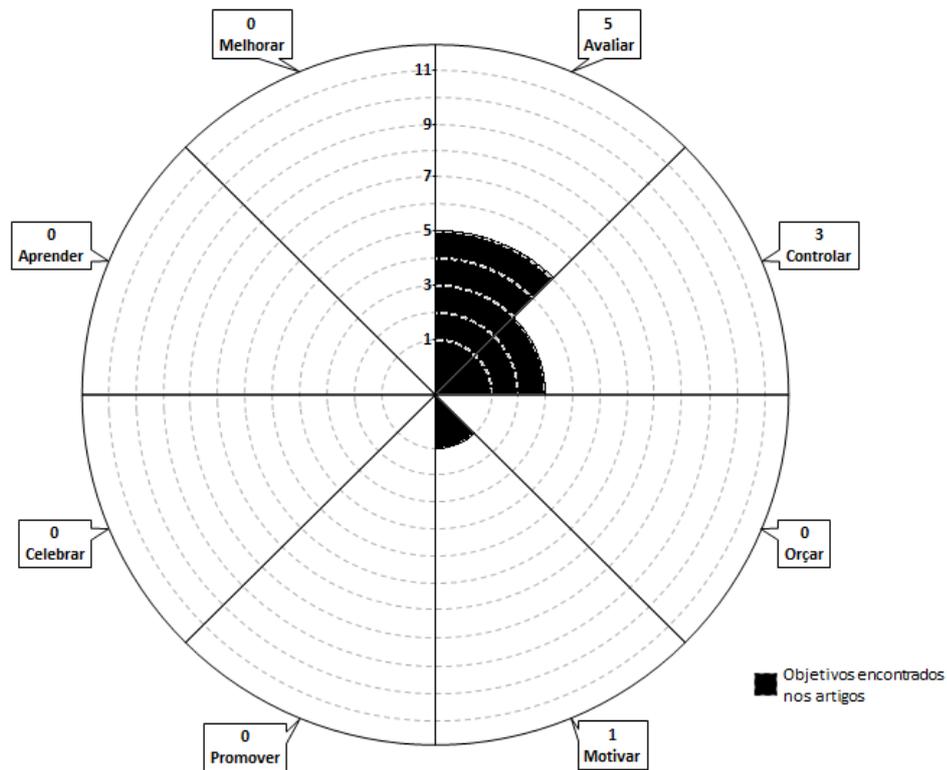


**Figura 10.** Stakeholders Identificados no PB

Fonte: dados da pesquisa.

Choong (2014) aponta a importância de definir os *stakeholders* da organização para que a informação correta seja transmitida. Chomchaiya e Esichaikul (2016) identificam as partes internas interessadas (administração, auditores, diretores financeiros, usuários do serviço e equipe de suporte do serviço) de compras públicas eletrônicas e, então, desenvolvem indicadores de desempenho com base na importância percebida pelos *stakeholders* internos.

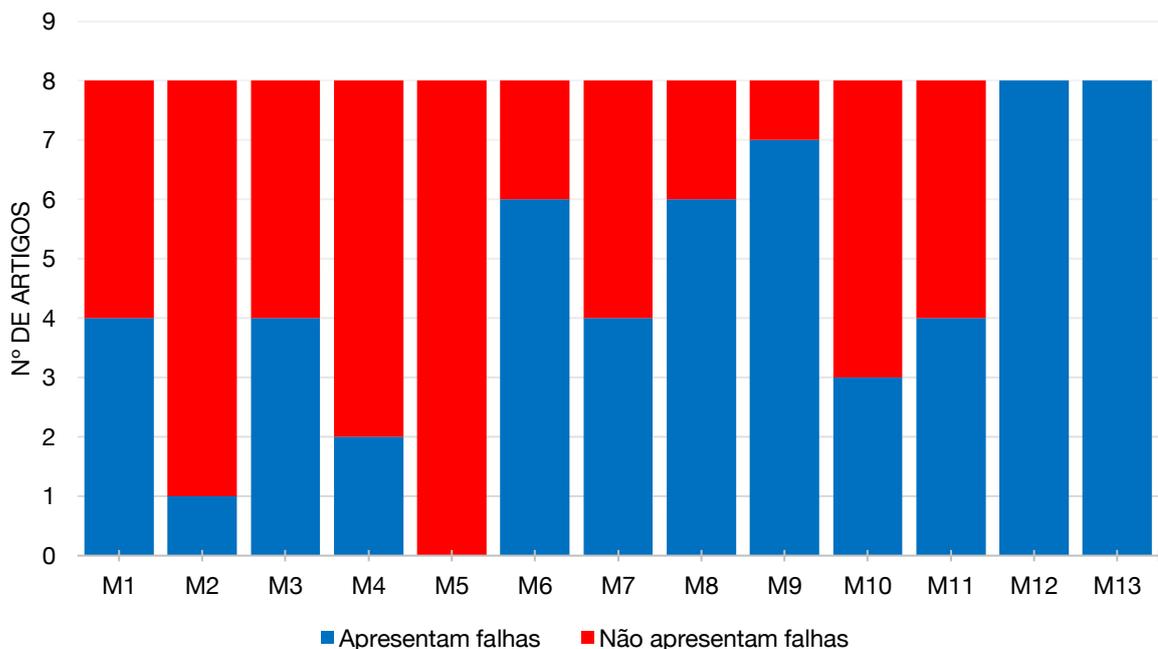
Na sequência, a Figura 11 apresenta os resultados encontrados sobre para quais propósitos, apontados por Behn (2003), os estudos utilizaram as medidas de desempenho. Dos oito propósitos citados pelo autor para os gestores públicos avaliarem o desempenho, identificou-se que somente os objetivos de avaliar, controlar e motivar foram considerados nos artigos. Cabe ressaltar que Behn (2003) trata dos propósitos específicos para avaliar o desempenho do setor público. Portanto, é de se estranhar que apenas três objetivos estejam contemplados na avaliação de desempenho de compras públicas.



**Figura 11.** Objetivos para Avaliação de Desempenho do Setor Público Identificados no PB

Fonte: dados da pesquisa.

Por fim, na última análise, identificaram-se as falhas que as medidas de desempenho apresentam, mostradas na Figura 12.



**Figura 12.** Falhas nas Métricas Encontradas

Fonte: dados da pesquisa.

Como destaque positivo estão as falhas M2, que apontam que as medidas de desempenho não são transpostas de outras empresas, e M5, pois não há foco excessivo em medidas financeiras, estando, nesse caso, as medidas bem equilibradas. Os problemas estão nas medidas de desempenho acabarem se tornando as metas. Isso é agravado pelo fato de as medidas não considerarem os objetivos organizacionais. Há, ainda, o fato de todos os estudos analisados apresentarem as falhas M12 e M13, que se referem às incertezas no início do projeto e à utilização de metas determinísticas, respectivamente.

O estudo das falhas reforça os problemas citados anteriormente, ou seja, as medidas de desempenho encontradas nos artigos do PB não estão alinhadas ao objetivo organizacional e não fornecem informações aos gestores para a tomada de decisão.

### Pauta de pesquisa

Na análise do fragmento de literatura foi possível identificar as lacunas de pesquisa que deram origem às possíveis contribuições, tanto para o meio acadêmico quanto para o meio gerencial (prático), apresentadas na Tabela 5.

**Tabela 6**

Contribuições de Pesquisa

Origem da Lacuna	Meio Acadêmico	Meio Gerencial (Prática)
Referencial Teórico	Demonstrar o ciclo de vida completo ao desenvolver ferramentas, não apenas o <i>design</i> .	Desenvolver sistemas de avaliação de desempenho completos para compras públicas.
Variável Básica	Aumentar a relação entre os autores da área, de forma a construir uma rede de pesquisa.  Cuidar para que os sistemas de avaliação de desempenho considerem os objetivos organizacionais.	Observar os estudos desenvolvidos pelos acadêmicos no desenvolvimento das ferramentas.  Considerar o <i>benchmarking</i> para a melhoria contínua, não somente para a comparação dos indicadores.
Variável Avançada	Desenvolver sistemas que considerem os interesses e possam informar os gestores.  Desenvolver formas para apresentar as medidas aos <i>stakeholders</i> .  Apresentar formas de desenvolver um sistema de avaliação de desempenho para compras públicas para casos específicos.  Construir um modelo/ferramenta que considere a melhoria contínua como objetivo.	Incluir o alinhamento estratégico nos modelos utilizados.  Utilizar os SAD de compras públicas para fornecer informações aos gestores para a tomada de decisão.  Buscar aprendizagem e melhoria contínua nos SADs, e não apenas avaliar e controlar os funcionários.  Identificar e buscar corrigir as falhas que os sistemas podem apresentar.

**Nota.** Fonte: dados da pesquisa.

Assim, identifica-se a necessidade de desenvolver uma ferramenta que contemple todos os aspectos teóricos de avaliação de desempenho, permitindo que os profissionais de compras públicas possam gerir esse processo da melhor maneira.

## Conclusões

A pesquisa teve por objetivo analisar um fragmento da literatura referente às compras públicas sob a perspectiva da avaliação de desempenho, visando a gerar conhecimento sobre o que tem sido realizado, quais lacunas demandam investigação, para, então, propor uma pauta para avanço nessa área. Para o processo de seleção, reflexão e análise do fragmento da literatura, utilizou-se o instrumento *ProKnow-C*, que conduziu à seleção de 16 artigos, que abordavam o fragmento da literatura desejado.

Por meio da análise bibliométrica das características básicas e avançadas delimitadas, evidenciam-se os conhecimentos gerados para serem socializados. Em linhas gerais, constatou-se: (i) a inexistência de autores renomados nessa área e de rede de colaboração; (ii) que os *stakeholders* são o foco que tem merecido atenção dos pesquisadores; (iii) que os modelos construídos não estão sendo capazes de oferecer subsídios à gestão de compras públicas; (iv) que o objetivo da AD tem sido avaliação, controle e motivação no âmbito das compras públicas; (v) destaque positivo para inexistência das falhas métricas de replicação de uma instituição para outra e foco excessivo em medidas financeiras; (vi) que, por outro lado, os modelos construídos para avaliar o desempenho não consideram os objetivos organizacionais e não fornecem informações aos gestores públicos; e (vii) forte preocupação com os *stakeholders*, tanto do contexto interno quanto do externo.

As descobertas feitas na pesquisa contribuem para o desenvolvimento da avaliação de desempenho de compras públicas e têm implicações para os profissionais e acadêmicos da área, conforme evidenciado pela Tabela 5, demonstrando o vasto caminho da AD de compras públicas para atender aos aportes teóricos da área de avaliação de desempenho. Entende-se que há espaço para os gestores utilizarem as medidas para alcançar o objetivo real da avaliação de desempenho, o qual, segundo Behn (2003), é melhorar o desempenho do governo.

A pesquisa está limitada ao campo amostral, pois somente foram consideradas as bases de dados selecionadas. Também se limita às publicações de língua inglesa, de 2000 a 2018, em periódicos internacionais, e à busca com palavras-chave selecionadas, bem como à percepção dos autores na seleção dos artigos. Outra limitação é referente ao processo de seleção dos artigos pelo *ProKnow-C*, pois foram realizadas somente duas etapas: seleção do PB e análise bibliométrica.

Assim, surgem, como sugestões de pesquisa futura, a ampliação do horizonte temporal, a inclusão de artigos escritos em outros idiomas e a investigação em outras bases de dados. Após a conclusão do estudo, questiona-se a possibilidade de exclusão do eixo “público” na seleção dos artigos, pois, segundo Rendon (2008), são utilizados os termos “*purchasing*” para tratar do tema no setor privado, e “*procurement*” ou “*acquisition*” no setor público.

Com relação ao instrumento de intervenção, é possível analisar o portfólio bibliográfico com base em outras características. Recomenda-se, ainda, a realização da análise sistêmica para completar as quatro etapas do *ProKnow-C*.

Considerando a influência da regulamentação nas compras públicas, sugere-se também o levantamento bibliográfico de países e/ou regiões específicas, para identificar como a legislação local afeta o ambiente da avaliação de desempenho de compras públicas. Por fim, com a pauta de pesquisa com base nas lacunas apontadas, sugere-se a criação de uma ferramenta, que permita a avaliação de desempenho de compras públicas que considere a percepção dos gestores e forneça informações aos *stakeholders*.

. Uma versão preliminar deste artigo foi apresentada no II Congresso Internacional de Desempenho do Setor Público (II CIDESP), realizado em Florianópolis, SC, em 2019.

## Referências

- Behn, R. D. (2003). Why measure performance? Different purposes require different measures. *Public Administration Review*, 63(5), 586-606. <https://doi.org/10.1111/1540-6210.00322>
- Bititci, U. S., Garengo, P., Ates, A., & Nudurupati, S. S. (2015). Value of maturity models in performance measurement. *International Journal of Production Research*, 53(10), 3062-3085. <https://doi.org/10.1080/00207543.2014.970709>
- Bititci, U., Garengo, P., Dörfler, V., & Nudurupati, S. (2012). Performance measurement: Challenges for tomorrow. *International Journal of Management Reviews*, 14(3), 305-327. <https://doi.org/10.1111/j.1468-2370.2011.00318.x>
- Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing and updating performance measurement systems. *International Journal of Operations & Production Management*, 20(7), 754-771. <https://doi.org/10.1108/01443570010330739>
- Brignall, S., & Modell, S. (2000). An institutional perspective on performance measurement and management in the 'new public sector'. *Management Accounting Research*, 11(3), 281-306. <https://doi.org/10.1006/mare.2000.0136>
- Camp, J. van, & Braet, J. (2016). Taxonomizing performance measurement systems' failures. *International Journal of Productivity and Performance Management*, 65(5), 672-693. <https://doi.org/10.1108/ijppm-03-2015-0054>
- Carneiro-da-Cunha, J. A., Hourneaux F., Jr, & Corrêa, H. L. (2016). Evolution and chronology of the organisational performance measurement field. *International Journal of Business Performance Management*, 17(2), 223-240. <https://doi.org/10.1504/ijbpm.2016.075553>
- Chomchaiya, S., & Esichaikul, V. (2016). Consolidated performance measurement framework for government e-procurement focusing on internal stakeholders.

*Information Technology & People*, 29(2), 354-380. <https://doi.org/10.1108/itp-12-2013-0210>

- Choong, K. K. (2014). Has this large number of performance measurement publications contributed to its better understanding? A systematic review for research and applications. *International Journal of Production Research*, 52(14), 4174-4197. <https://doi.org/10.1080/00207543.2013.866285>
- Ensslin, S. R., Ensslin, L., Matos, L. dos S., Dutra, A., & Ripoll-Feliu, V. M. (2015). Research opportunities in performance measurement in public utilities regulation. *International Journal of Productivity and Performance Management*, 64(7), 994-1017. <https://doi.org/10.1108/ijppm-05-2014-0067>
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S., Marr, B., & Neely, A. (2007). Towards a definition of a business performance measurement system. *International Journal of Operations & Production Management*, 27(8), 784-801. <https://doi.org/10.1108/01443570710763778>
- Galera, A. N., Rodríguez, D. O., & Hernandez, A. M. L. (2008). Identifying barriers to the application of standardized performance indicators in local government. *Public Management Review*, 10(2), 241-262. <https://doi.org/10.1080/14719030801928706>
- Gardenal, F. (2013). A model to measure e-procurement impacts on organizational performance. *Journal of Public Procurement*, 13(2), 215-242. <https://doi.org/10.1108/jopp-13-02-2013-b003>
- Ghalayini, A. M., & Noble, J. S. (1996). The changing basis of performance measurement. *International Journal of Operations & Production Management*, 16(8), 63-80. <https://doi.org/10.1108/01443579610125787>
- Goh, S. C., Elliott, C., & Richards, G. (2015). Performance management in Canadian public organizations: Findings of a multi-case study. *International Journal of Productivity and Performance Management*, 64(2), 157-174. <https://doi.org/10.1108/ijppm-10-2013-0170>
- Kim, S. W., Olivares, M., & Weintraub, G. Y. (2014). Measuring the performance of large-scale combinatorial auctions: A structural estimation approach. *Management Science*, 60(5), 1180-1201. <https://doi.org/10.1287/mnsc.2013.1814>
- Lebas, M. J. (1995). Performance measurement and performance management. *International Journal of Production Economics*, 41(1-3), 23-35. [https://doi.org/10.1016/0925-5273\(95\)00081-X](https://doi.org/10.1016/0925-5273(95)00081-X)
- Mardale, F. E. (2015). Performance in public procurement. *Land Forces Academy Review*, 20(4), 476-477.

- Matos, L. dos S., Ensslin, S. R. & Ensslin, L. (2019). A review on the performance measurement systems life cycle elements. *Lex Localis - Journal of Local Self-Government*, 17(4), 939-959. [https://doi.org/10.4335/17.4.939-959\(2019\)](https://doi.org/10.4335/17.4.939-959(2019))
- Matos, L. dos S., Valmorbidia, S. M. I., Martins, V. A., & Ensslin, S. R. (2019). Development of performance evaluation theme: A systematic analysis of the literature. *Contextus - Revista Contemporânea de Economia e Gestão*, 17(2), 63-97. <https://doi.org/10.19094/contextus.v17i2.40146>
- Melnyk, S. A., Bititci, U., Platts, K., Tobias, J., & Andersen, B. (2014). Is performance measurement and management fit for the future? *Management Accounting Research*, 25(2), 173-186. <https://doi.org/10.1016/j.mar.2013.07.007>
- Neely, A. (1999). The performance measurement revolution: Why now and what next? *International Journal of Operations & Production Management*, 19(2), 205-228. <https://doi.org/10.1108/01443579910247437>
- Neely, A., Gregory, M., & Platts, K. (1995). Performance measurement system design: A literature review and research agenda. *International Journal of Operations & Production Management*, 15(4), 80-116. <https://doi.org/10.1108/01443579510083622>
- Nudurupati, S. S., Bititci, U. S., Kumar, V., & Chan, F. T. S. (2011). State of the art literature review on performance measurement. *Computers & Industrial Engineering*, 60(2), 279-290. <https://doi.org/10.1016/j.cie.2010.11.010>
- Patrucco, A. S., Luzzini, D., & Ronchi, S. (2016). Evaluating the effectiveness of public procurement performance management systems in local governments. *Local Government Studies*, 42(5), 739-761. <https://doi.org/10.1080/03003930.2016.1181059>
- Ren, Z., Kwaw, P., & Yang, F. (2012). Ghana's public procurement reform and the continuous use of the traditional procurement system: The way forward. *Built Environment Project and Asset Management*, 2(1), 56-69. <https://doi.org/10.1108/20441241211235053>
- Rendon, R. G. (2008). Procurement process maturity: Key to performance measurement. *Journal of Public Procurement*, 8(2), 200-214. <https://doi.org/10.1108/jopp-08-02-2008-b003>
- Reuter, C., Goebel, P., & Foerstl, K. (2012). The impact of stakeholder orientation on sustainability and cost prevalence in supplier selection decisions. *Journal of Purchasing and Supply Management*, 18(4), 270-281. <https://doi.org/10.1016/j.pursup.2012.06.004>
- Richardson, R. J. (1999). *Pesquisa social: Métodos e técnicas* (3a ed.). São Paulo: Atlas.

- Rotchanakitumnuai, S. (2013). Assessment of e-procurement auction with a balanced scorecard. *International Journal of Physical Distribution & Logistics Management*, 43(1), 39-53. <https://doi.org/10.1108/09600031311293246>
- Schapper, P. R., Malta, J. N. V., & Gilbert, D. L. (2017). Analytical framework for the management and reform of public procurement. In K. V. Thai (Ed.), *International handbook of public procurement* (pp. 87-104). New York: Routledge. <https://doi.org/10.4324/9781315092539-5>
- Schiele, J. J. (2005). A tool for assessing the value contributed by public purchasing departments throughout various stages of competitive acquisition processes for consulting services. *Journal of Public Procurement*, 5(1), 73-85. <https://doi.org/10.1108/jopp-05-01-2005-b004>
- Srivastava, S. K. & Agrahari, A. (2017). Benchmarking approach to improve the public procurement process. *Economic and Political Weekly*, 52(20), 58-67.
- Thiel, G. G., Ensslin, S. R., & Ensslin, L. (2017). Street lighting management and performance evaluation: Opportunities and challenges. *Lex Localis - Journal of Local Self-Government*, 15(2), 303-328. [https://doi.org/10.4335/15.2.303-328\(2017\)](https://doi.org/10.4335/15.2.303-328(2017))
- Tillema, S., Mimba, N. P. S. H., & Helden, G. J. van (2010). Understanding the changing role of public sector performance measurement in less developed countries. *Public Administration and Development*, 30(3), 203-214. <https://doi.org/10.1002/pad.561>
- Valmorbida, S. M. I., Ensslin, S. P. R., Ensslin, L., & Ripoll-Feliu, V. M. (2016). Rankings universitários mundiais: Que dizem os estudos internacionais? *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 14(2), 5-29. <https://doi.org/10.15366/reice2016.14.2.001>