A MENSURAÇÃO DA TRANSFORMAÇÃO DIGITAL EM GESTÃO ESTRATÉGICA: UMA REVISÃO DA VALIDADE DE CONSTRUTO

THE MEASUREMENT OF DIGITAL TRANSFORMATION IN STRATEGIC MANAGEMENT: A CONSTRUCT VALIDITY REVIEW

JOELMA MARIA PEREIRA SIMPLICIO
UNINOVE – UNIVERSIDADE NOVE DE JULHO

FERNANDO ANTONIO RIBEIRO SERRA
UNINOVE – UNIVERSIDADE NOVE DE JULHO

LEANDRO SIMPLICIO SILVA
UNINOVE – UNIVERSIDADE NOVE DE JULHO

Comunicação:
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

Agradecimento à orgão de fomento:
O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

A MENSURAÇÃO DA TRANSFORMAÇÃO DIGITAL EM GESTÃO ESTRATÉGICA: UMA REVISÃO DA VALIDADE DE CONSTRUTO

Objetivo do estudo
Avaliar se os instrumentos utilizados pela literatura empírica para mensurar a transformação digital em gestão estratégica atendem aos padrões de validade de construto estabelecidos pela literatura metodológica, por meio de uma auditoria sistemática das medidas empregadas no campo.

Relevância/originalidade
A transformação digital é um dos construtos mais estudados e menos consistentemente mensurados da gestão estratégica. O estudo submete suas medidas ao mesmo escrutínio de validade de construto exigido de outros construtos, lacuna ainda não endereçada de forma sistemática.

Metodologia/abordagem
Auditoria em duas fases: síntese cruzada de onze revisões de referência, da qual se derivam cinco critérios de adequação da mensuração (C1–C5), seguida da auditoria de 109 estudos empíricos frente a esses critérios, classificados por família de instrumento.

Principais resultados
Nenhum dos noventa instrumentos empiricamente aplicados satisfaz simultaneamente os cinco critérios. As violações seguem um padrão estrutural por família: mineração textual e proxies capturam visibilidade; índices, escalas e modelos de maturidade capturam intensidade percebida, não o construto teórico.

Contribuições teóricas/metodológicas
Deriva cinco critérios explícitos de adequação da mensuração, oferecendo um protocolo replicável para avaliar instrumentos existentes e desenhar novos. Torna explícito que parte do conhecimento acumulado pode refletir o que as medidas capturam, e não o fenômeno.

Contribuições sociais/para a gestão
Escores de transformação digital orientam decisões de investidores, gestores e reguladores. Ao evidenciar que medidas dominantes capturam visibilidade, não alinhamento estratégico, o estudo alerta que decisões calibradas por esses indicadores podem mirar o alvo errado.

Palavras-chave: Mensuração da transformação digital, mensuração de construtos, Gestão estratégica, Validade de construto, Modelos formativos

THE MEASUREMENT OF DIGITAL TRANSFORMATION IN STRATEGIC MANAGEMENT: A CONSTRUCT VALIDITY REVIEW

Study purpose
To assess whether the instruments used in the empirical literature to measure digital transformation in strategic management meet the construct-validity standards established by the methodological literature, through a systematic audit of the measures employed in the field.

Relevance / originality
Digital transformation is among the most studied and least consistently measured constructs in strategic management. The study subjects its measures to the same construct-validity scrutiny required of other constructs, a gap not yet addressed systematically in the field.

Methodology / approach
A two-phase audit: a cross-synthesis of eleven reference reviews, from which five measurement-adequacy criteria (C1–C5) are derived, followed by an audit of 109 empirical studies against these criteria, classified by instrument family.

Main results
None of the ninety empirically applied instruments satisfies all five criteria simultaneously. Violations follow a structural pattern by family: text mining and proxies capture visibility; indices, scales, and maturity models capture perceived intensity, not the theoretical construct.

Theoretical / methodological contributions
Derives five explicit measurement-adequacy criteria, offering a replicable protocol to evaluate existing instruments and design new ones. It makes explicit that part of the accumulated knowledge may reflect what the measures capture, rather than the phenomenon itself.

Social / management contributions
Digital-transformation scores guide decisions by investors, managers, and regulators. By showing that dominant measures capture visibility rather than strategic alignment, the study warns that decisions calibrated on these indicators may aim at the wrong target.

Keywords: Digital transformation measurement, Construct measurement, Strategic management, Construct validity, Formative models

A MENSURAÇÃO DA TRANSFORMAÇÃO DIGITAL EM GESTÃO ESTRATÉGICA: UMA REVISÃO DA VALIDADE DE CONSTRUTO

1. Introdução

A transformação digital (TD) tornou-se um dos fenômenos organizacionais centrais da última década e um dos domínios de pesquisa mais ativos na gestão estratégica e nos sistemas de informação contemporâneos (Hanelt, Bohnsack, Marz, & Antunes Marante, 2021; Vial, 2019). A literatura empírica acumulou várias centenas de estudos que estimam a relação entre TD e resultados em nível da firma, incluindo desempenho, posicionamento competitivo e mudança organizacional (Hanelt et al., 2021; Verhoef et al., 2021). O construto é mobilizado em pesquisas sobre sistemas de informação, mudança organizacional e gestão estratégica, com diversas revisões de literatura recentes documentando sua expansão e influência (Hausberg, Liere-Netheler, Packmohr, Pakura, & Vogelsang, 2019; Nadkarni & Prügl, 2021; Reis & Melão, 2023).

O construto, entretanto, é objeto de disputas teóricas. A literatura oferece diversas definições coexistentes, tratando a TD ora como um processo temporal de adoção tecnológica em fases (Reis et al., 2018; Vial, 2019), ora como uma mudança organizacional desencadeada pela difusão de tecnologias digitais (Hanelt et al., 2021; Nadkarni & Prügl, 2021), ora como uma redefinição do modelo de negócios em torno da criação de valor digital (Schallmo et al., 2017; Verhoef et al., 2021). Cada definição implica diferentes observáveis, fronteiras conceituais e requisitos de operacionalização (Boyd et al., 2013; Suddaby, 2010). Paralelamente, a literatura metodológica em gestão estratégica vem estabelecendo, há quase quatro décadas, critérios para a mensuração adequada de construtos, incluindo clareza conceitual, validade de conteúdo, distinção entre mensuração reflexiva e formativa, dimensionalidade e nível apropriado de análise (Boyd et al., 2013; Edwards, 2001; Edwards & Bagozzi, 2000; MacKenzie, Podsakoff, & Podsakoff, 2011; Suddaby, 2010; Venkatraman & Grant, 1986). Diversas revisões sobre TD observam que o campo carece de medidas consistentes e que a comparabilidade entre estudos permanece limitada (Hanelt et al., 2021; Hausberg et al., 2019; Jamaludin & Basari, 2022; Jedynak et al., 2021; Reis & Melão, 2023; Verhoef et al., 2021). Algumas dessas revisões identificam ainda pressupostos incorporados às medidas dominantes, incluindo monotonicidade (Jedynak et al., 2021) e a ausência de consideração de efeitos negativos da TD (Feliciano-Cestero et al., 2023).

Apesar dessas evidências, a literatura sobre TD ainda não foi objeto de uma auditoria estruturada que avalie seus instrumentos operacionais à luz dos padrões de validade de construto amplamente estabelecidos na área. Evidências bibliométricas sugerem que a literatura empírica tem convergido para uma narrativa de otimização, na qual a TD é tratada predominantemente como um fenômeno gerador de valor (Aurélio, 2024). Ainda não está claro, entretanto, se essa mesma convergência se manifesta nos instrumentos utilizados para operacionalizar o construto. A lacuna abordada por este estudo é, portanto, metodológica e estrutural: verificar se as medidas dominantes de TD atendem aos requisitos de validade que a própria disciplina estabelece para a mensuração de construtos.

A partir dessa lacuna, este estudo busca responder à seguinte questão de pesquisa: os instrumentos dominantes de mensuração da TD satisfazem os requisitos de validade de construto amplamente reconhecidos na literatura metodológica de gestão estratégica? Para responder a essa questão, o estudo adota um desenho de pesquisa em duas fases. A Fase 1 extrai

o conteúdo substantivo de onze revisões de literatura de referência sobre TD e deriva, a partir dessa síntese, cinco critérios de adequação da mensuração que qualquer instrumento válido deveria satisfazer. A Fase 2 audita um corpus de aproximadamente cem artigos empíricos quantitativos com base nesses critérios.

Os resultados mostram que nenhum instrumento presente no corpus satisfaz simultaneamente os cinco critérios e que as violações se agrupam de forma estruturalmente consistente entre as diferentes famílias de instrumentos. Instrumentos baseados em mineração de texto e proxies de bases de dados capturam predominantemente o gatilho tecnológico da TD; instrumentos construídos a partir de índices compostos, escalas e modelos de maturidade capturam aspectos da transformação organizacional, mas operam com especificações reflexivas que comprometem a independência dimensional requerida pelo construto; e todos os instrumentos analisados registram efeitos apenas no nível da firma, deixando de contemplar consequências para ecossistemas, mercado de trabalho e sociedade.

A contribuição deste estudo é dupla. Primeiro, traduz os padrões de validade de construto consolidados na literatura metodológica de gestão estratégica em critérios operacionais específicos para a avaliação da mensuração da TD. Segundo, demonstra que parte das evidências empíricas acumuladas sobre a relação entre TD e desempenho organizacional requer reinterpretação à luz dos critérios de validade de construto adotados pela própria disciplina.

2. Referencial Teórico

2.1 Transformação digital como construto

A TD foi formalizada na gestão estratégica e nos sistemas de informação por meio de diversas formulações que compartilham o foco na mudança organizacional associada às tecnologias digitais, mas diferem quanto ao conteúdo substantivo do construto (Hanelt et al., 2021; Nadkarni & Prügl, 2021; Vial, 2019). Uma primeira tradição caracteriza a TD como um processo composto por fases sequenciais de digitização, digitalização e transformação, em que a adoção tecnológica constitui o antecedente e a reconfiguração organizacional representa a trajetória do processo (Reis et al., 2018; Verhoef et al., 2021; Vial, 2019). Uma segunda tradição estabelece uma distinção categórica entre a tecnologia como gatilho e a mudança organizacional como conteúdo, definindo a TD como "mudança organizacional desencadeada e moldada pela ampla difusão de tecnologias digitais" (Hanelt et al., 2021, p. 1160) ou como "transformação organizacional conduzida por atores e desencadeada pela adoção de disrupções digitais impulsionadas por tecnologia" (Nadkarni & Prügl, 2021, p. 240). Uma terceira tradição centra a transformação no modelo de negócios como unidade de análise, definindo o construto como "uma mudança na forma como uma empresa emprega tecnologias digitais para desenvolver um novo modelo de negócios digital que contribui para criar e apropriar mais valor para a firma" (Verhoef et al., 2021, p. 889) ou por meio de uma reconfiguração em cinco fases dos mecanismos de criação, entrega e captura de valor (Schallmo, Williams, & Boardman, 2017).

Essas formulações não são intercambiáveis. Elas induzem diferentes observáveis, sugerem diferentes unidades de análise e são sustentadas por conjuntos teóricos parcialmente sobrepostos (Bharadwaj et al., 2013; Boyd et al., 2013; Matt et al., 2015; Sebastian et al., 2017; Suddaby, 2010). Consequentemente, a escolha da definição adotada possui implicações diretas para a operacionalização do construto e para a avaliação da adequação de suas medidas (Boyd et al., 2013; MacKenzie, Podsakoff, & Podsakoff, 2011).

Entre as formulações existentes, a proposta por Hanelt et al. (2021) tornou-se uma das mais influentes na literatura recente e apresenta uma vantagem analítica relevante para os objetivos deste estudo. Ao distinguir explicitamente o gatilho da TD, representado pela difusão

de tecnologias digitais, de seu conteúdo, representado pelas mudanças organizacionais associadas a essa difusão, a definição fornece um critério mais preciso para avaliar a adequação da mensuração do construto. Assim, a definição adotada neste estudo é a de Hanelt et al. (2021), enquanto as formulações alternativas são interpretadas como descrições de como a TD se desenvolve ao longo do tempo (a trajetória da tradição processual) e daquilo que pode constituir seu conteúdo substantivo (a reconfiguração do modelo de negócios enfatizada pela terceira tradição).

2.2 Mensuração de construtos no campo da gestão estratégica

A mensuração na pesquisa em gestão estratégica tem sido reconhecida como uma fonte recorrente de limitações metodológicas há pelo menos quatro décadas, e sucessivas revisões documentaram avanços limitados apesar do aparato metodológico disponível para a área (Boyd et al., 2013; Boyd et al., 2005; Hitt et al., 2004; Venkatraman & Grant, 1986). Entre os problemas mais frequentes estão o uso de indicadores únicos para construtos complexos, a divulgação insuficiente de evidências de confiabilidade, ambiguidades quanto ao nível de análise e a tendência de tratar a mensuração como uma etapa técnica subordinada à teoria, em vez de uma escolha teórica em si mesma (Boyd et al., 2013). Em resposta, a literatura metodológica consolidou seis componentes centrais da validade de construto em estudos quantitativos: validade de conteúdo, dimensionalidade, confiabilidade, validade convergente, validade discriminante e validade preditiva (Boyd et al., 2013; Venkatraman & Grant, 1986).

A clareza conceitual constitui uma condição necessária para o uso cumulativo de construtos na pesquisa empírica. Quando o conteúdo conceitual, as fronteiras do construto e a relação entre construto e indicadores não são claramente especificadas, o mesmo termo pode ser utilizado para representar fenômenos distintos, comprometendo a comparabilidade dos achados (Suddaby, 2010). Nessas circunstâncias, resultados produzidos por estudos que utilizam a mesma terminologia podem, na realidade, referir-se a objetos empíricos diferentes. A consequência é a acumulação de evidências cuja integração e interpretação tornam-se problemáticas. Além disso, erros de especificação no nível da mensuração produzem vieses inferenciais que não podem ser corrigidos por sofisticação estatística (Edwards, 2001; MacKenzie et al., 2011).

Para construtos multidimensionais, uma das decisões mais relevantes refere-se à escolha entre modelos de mensuração reflexivos e formativos. Em uma especificação reflexiva, o construto latente causa seus indicadores, que representam manifestações intercambiáveis de uma mesma entidade e, portanto, devem apresentar elevada covariação (Edwards & Bagozzi, 2000; Jarvis, MacKenzie, & Podsakoff, 2003). Em uma especificação formativa, os indicadores constituem conjuntamente o construto, podem variar independentemente entre si e não são mutuamente substituíveis (Diamantopoulos & Winklhofer, 2001; Jarvis et al., 2003).

A escolha entre essas especificações deve decorrer da natureza teórica do fenômeno investigado (Edwards & Bagozzi, 2000; Jarvis et al., 2003). Tratar um construto formativo como reflexivo elimina a estrutura dimensional definida pela teoria como constitutiva do construto (Boyd et al., 2013; Howell et al., 2007; MacKenzie et al., 2011; Podsakoff, Shen, & Podsakoff, 2006). Na gestão estratégica, essa prática tem sido recorrente, resultando na mensuração de construtos amplos e multidimensionais por meio de indicadores únicos ou de escalas cuja arquitetura dimensional raramente é examinada diretamente (Boyd et al., 2013; Podsakoff et al., 2006).

2.3 Justificativa para uma auditoria da mensuração

A aplicação dos princípios de validade de construto a uma literatura específica depende das características de desenvolvimento desse campo de pesquisa (Boyd et al., 2013; Venkatraman & Grant, 1986). Embora esses princípios sejam amplamente consolidados na gestão estratégica, sua relevância empírica varia conforme os padrões de conceituação e operacionalização adotados em cada domínio. No caso da TD, diferentes evidências sugerem a necessidade de uma avaliação sistemática dos instrumentos utilizados para mensurar o construto.

Uma primeira evidência decorre de análises bibliométricas que apontam para a convergência da literatura empírica em torno de uma narrativa de otimização, na qual a TD é predominantemente associada à criação de valor e à melhoria de desempenho (Aurélio, 2024). Quando essa convergência também se manifesta nos indicadores utilizados para representar o construto, parte das conclusões acumuladas pelo campo pode refletir escolhas de mensuração tanto quanto características do fenômeno investigado (Boyd et al., 2013; MacKenzie et al., 2011). Essa possibilidade justifica o exame direto dos instrumentos empregados pela literatura.

Uma segunda evidência relaciona-se às características de campos em rápida expansão. Nessas circunstâncias, práticas de mensuração tendem a se institucionalizar antes de serem submetidas a avaliações sistemáticas, fazendo com que proxies inicialmente adotadas por conveniência sejam reproduzidas ao longo do tempo (Boyd et al., 2013; Pfeffer, 1993). Padrões semelhantes já foram observados em outros construtos da gestão estratégica e do empreendedorismo (Crook et al., 2010; Godfrey & Hill, 1995). A literatura sobre TD apresenta características compatíveis com essa dinâmica.

Uma terceira evidência provém das revisões de literatura disponíveis. Esses estudos documentam dificuldades persistentes relacionadas à mensuração da TD e apontam limitações quanto à consistência das medidas, à delimitação conceitual do construto e à comparabilidade dos resultados empíricos (Hanelt et al., 2021; Hausberg et al., 2019; Jamaludin & Basari, 2022; Jedynak et al., 2021; Reis & Melão, 2023; Verhoef et al., 2021). Algumas revisões identificam ainda pressupostos incorporados às medidas dominantes, incluindo monotonicidade (a premissa de que a relação entre TD e desempenho é sempre positiva e crescente), efeitos negativos pouco considerados e uma ênfase predominante no nível da firma (Feliciano-Cestero et al., 2023; Hausberg et al., 2019; Jedynak et al., 2021; Nadkarni & Prügl, 2021). Em conjunto, essas evidências indicam a necessidade de uma auditoria estruturada dos instrumentos de mensuração utilizados na pesquisa sobre TD.

3. Método

Este estudo examina criticamente os instrumentos utilizados para mensurar a TD na literatura empírica. Em consonância com as recomendações recentes para revisões focadas em métodos (Block et al., 2025), o objetivo é avaliar a aderência desses instrumentos aos princípios de validade de construto estabelecidos na gestão estratégica (Boyd et al., 2013). Para isso, foi desenvolvido um desenho de pesquisa em duas fases sequenciais.

A Fase 1 realiza uma síntese cruzada estruturada das principais revisões de literatura sobre TD e deriva os critérios de adequação da mensuração utilizados neste estudo. A Fase 2 audita um corpus de artigos empíricos que operacionalizam a TD com base nesses critérios e identifica os padrões de violações observados entre os diferentes instrumentos.

As revisões foram selecionadas por meio da combinação de buscas sistemáticas nas bases Web of Science (Core Collection) e Scopus com a recuperação direcionada de revisões frequentemente citadas sobre TD nas áreas de gestão estratégica e sistemas de informação. As buscas utilizaram combinações dos termos "*digital transformation*", "*digitalization*", "*digitization*", "*literature review*", "*systematic review*", "*review*" e "*survey*", aplicados a

títulos, resumos e palavras-chave. Foram aplicados três critérios de inclusão: escopo explicitamente definido como revisão de literatura, publicação em periódico com revisão por pares e centralidade demonstrada na literatura, medida por impacto de citações ou recorrência em revisões subsequentes. Onze revisões atenderam simultaneamente aos três critérios e compõem o corpus da Fase 1. Essas revisões abrangem o período de 2017 a 2023 e incluem diferentes abordagens metodológicas, como revisões sistemáticas, revisões de escopo, meta-revisões e revisões baseadas em redes de citações. A Tabela 1 apresenta o corpus analisado.

Tabela 1.

Corpus de revisões de literatura utilizadas na Fase 1.

ID	Referência (periódico)	Tipo e corpus	Principal contribuição	Justificativa para inclusão
R01	Vial (2019), JSIS	Síntese indutiva de trabalhos existentes (282 estudos de SI)	Definição por decomposição semântica; blocos constituintes do processo	Marco definicional mais citado; referência central para a tradição processual
R02	Verhoef et al. (2021), JBR	Revisão de escopo multidisciplinar (84 artigos)	Três fases hierárquicas; identifica a mensuração e a progressão entre fases como lacuna aberta	Frequentemente citado em diferentes famílias; explicita a ausência de medidas válidas
R03	Nadkarni and Prügl (2021), MRQ	Revisão sistemática (58 + 6 estudos adjacentes)	Duas dimensões agregadas (tecnologia e ator) com nove temas; identifica lacuna na gestão intermediária	Distinção rigorosa entre tecnologia e mudança; delimitação explícita das fronteiras da firma como escolha analítica
R04	Jamaludin and Basari (2022), SMRJ	SALSA (20 artigos, estratégia de TD)	Diagnostica o campo como imaturo, fragmentado e carente de mensuração consistente	Evidência direta sobre o estado da mensuração na literatura de estratégia
R05	Reis et al. (2018), AISC	Revisão bibliométrica e qualitativa (206 artigos)	Definição tripartite (tecnologia, organização e aspectos sociais); mapeamento das fases da transformação digital	Inventário dimensional fundamental adotado por revisões posteriores
R06	Schallmo et al. (2017), IJIM	Estrutura conceitual com ilustração de caso	Roteiro em cinco fases para a transformação de modelos de negócios digitais	Referência de base para o modelo de maturidade de negócios
R07	Jedynak et al. (2021), JOCM	Revisão sistemática com literatura cinzenta (234 publicações acadêmicas + 27 de consultoria)	Diagnostica a monotonicidade de forma explícita; identifica lacuna relacionada aos efeitos negativos	Crítica metodológica mais desenvolvida aos pressupostos de mensuração presentes no campo
R08	Reis and Melão (2023), Heliyon	Meta-revisão (PRISMA, 34 revisões sistemáticas de literatura)	Seis dimensões e dezessete categorias; explicita o paradoxo da transformação digital	Primeira revisão de segunda ordem do campo; consolida o inventário dimensional entre revisões
R09	Feliciano-Cestero et al. (2023), JBR	Revisão sistemática (123 artigos, transformação digital e internacionalização)	Argumenta que os efeitos negativos da transformação digital são subexplorados	Uma das poucas revisões que problematiza explicitamente o pressuposto da monotonicidade

R10	Hanelt et al. (2021), JMS	Revisão sistemática (279 artigos de periódicos de primeira linha)	Condições, mecanismos e resultados enquadrados por ecossistemas turbulentos	Fonte da definição integradora adotada na Seção 2.1
R11	Hausberg et al. (2019), JBE	Revisão sistemática da literatura com rede de citações (1.876 artigos)	Nove correntes de pesquisa; analisa os efeitos sobre sociedade, controle e poder	Maior corpus analisado; única revisão com cluster explícito sobre sociedade e poder

A síntese cruzada integrou as informações extraídas das onze revisões com o objetivo de identificar padrões recorrentes na conceituação e na mensuração da TD. A análise produziu três resultados centrais: (i) uma tipologia de famílias definicionais, obtida pelo agrupamento das definições segundo o conteúdo substantivo atribuído ao construto; (ii) um inventário dimensional consolidado, derivado do alinhamento das estruturas dimensionais apresentadas pelas revisões; e (iii) uma codificação dos pressupostos implícitos associados à mensuração, registrada em termos de presença, ausência, ambiguidade ou problematização em cada revisão.

Esses resultados constituíram a base para a derivação dos critérios C1–C5 de adequação da mensuração, como apresentado na Tabela 2 a seguir. As famílias definicionais forneceram subsídios para avaliar a correspondência entre definição e operacionalização; o inventário dimensional orientou a análise da cobertura do construto; e a codificação dos pressupostos permitiu identificar aspectos sistematicamente negligenciados pelos instrumentos existentes. Como resultado complementar, foi também elaborada uma rede agregada das referências mobilizadas pelas revisões, utilizada para identificar as principais fontes fundacionais do campo. Os resultados da síntese cruzada e os critérios derivados são apresentados na Seção 4.1.

Tabela 2.

Critérios de adequação da mensuração (C1–C5) e suas dimensões de auditoria.

Critério	Dimensão de auditoria	Evidências examinadas	Categorias de codificação	Interpretação diagnóstica
C1	Compromisso explícito com a definição adotada	Família definicional adotada (processo, mudança desencadeada pela tecnologia ou redefinição do modelo de negócios); referência teórica utilizada; coerência entre a definição declarada e o instrumento empregado	FAM-A / FAM-B / FAM-C / NÃO DECLARADA; coerência: SIM / PARCIAL / NÃO	Uma família não declarada equivale a um compromisso implícito com a família que a fonte de dados torna operacionalmente viável
C2	Foco no gatilho versus conteúdo	Tipo de dado utilizado (mineração textual de relatórios, patentes, investimentos em TI, survey de mudança organizacional, modelo de maturidade digital, escala psicométrica); o que o instrumento efetivamente mensura	GATILHO / CONTEÚDO / HÍBRIDO; tipo de dado (rótulo)	Um instrumento focado no gatilho, mas apresentado como medida substantiva de transformação digital, indica desalinhamento entre definição teórica e operacionalização
C3	Cobertura dimensional e tratamento da independência	Quais das seis dimensões centrais o instrumento contempla (tecnológica, estratégica, organizacional,	Cobertura: 1 a 6; especificação: REFLEXIVA / FORMATIVA /	Cobertura inferior a quatro dimensões associada a especificação reflexiva

		processos, ecossistema e pessoas); especificação reflexiva ou formativa; presença de validação entre dimensões	COMPÓSITA / NÃO ESPECIFICADA	indica colapso da estrutura formativa do construto
C4	Discriminação entre não adoção estratégica e falha de implementação	Presença de proxy para intenção estratégica; tratamento de valores baixos no índice; discussão da não adoção como escolha estratégica	DISCRIMINA / NÃO DISCRIMINA; tratamento dos zeros: VAZIO / FALHA / ESCOLHA / AMBÍGUO	A incapacidade de discriminar implica que o coeficiente estimado incorpora, por construção, a hipótese de que mais transformação digital é sempre melhor
C5	Nível de registro e externalidades	Nível em que a medida é construída; efeitos teoricamente relevantes não capturados pelo instrumento (trabalho, fornecedores, ecossistema, sociedade)	Nível: FIRMA / SETOR / ECOSSISTEMA / MULTINÍVEL; externalidades: lista nominal	O registro apenas no nível da firma, diante de efeitos teorizados além da firma, indica externalização da heterogeneidade distributiva pelo próprio desenho do instrumento

Nota: Cada critério corresponde a uma dimensão analítica com indicadores observáveis, regras de codificação e interpretação diagnóstica dos resultados. As dimensões associadas ao C3 derivam do inventário dimensional consolidado da Fase 1, enquanto as associadas aos critérios C4 e C5 derivam da codificação dos pressupostos identificados nas revisões analisadas. Os critérios são aplicados aos instrumentos descritos nos artigos, e não aos resultados estatísticos reportados, de modo que a auditoria avalia a adequação da mensuração e não a robustez das estimações realizadas.

A Fase 2 consiste na auditoria dos artigos empíricos quantitativos que operacionalizam a TD no nível da firma. Para isso, os cinco critérios derivados na Fase 1 (C1-C5) são operacionalizados como dimensões de auditoria aplicadas a cada instrumento identificado no corpus. A auditoria registra: (i) a família definicional implícita no instrumento e se ela é explicitamente declarada (C1); (ii) se a medida captura o gatilho ou o conteúdo da transformação digital (C2); (iii) quais das seis dimensões centrais são contempladas pelo instrumento e como sua independência é tratada (C3); (iv) se o instrumento distingue a não adoção estratégica de uma falha de implementação (C4); e (v) o nível em que o construto é registrado e quais efeitos extrafirma são externalizados pela medida (C5). O objetivo é identificar padrões recorrentes de adequação e violação dos critérios de mensuração na literatura empírica sobre transformação digital.

O corpus inicial da auditoria foi constituído por uma base unificada de 151 registros obtidos por meio de buscas paralelas nas bases Web of Science (Core Collection) e Scopus. As buscas combinaram termos relacionados à transformação digital *("digital transformation" e "digitalization"*) com termos associados à sua mensuração e operacionalização, incluindo expressões como "*digital transformation index*", "*digital maturity*", "*digital maturity model*", "*measurement", "proxy*" e "*digital capability*", aplicadas aos campos de título, resumo e palavras-chave. Após a consolidação das bases, foram removidas duplicatas exatas, identificadas por DOI ou por título normalizado. Durante o desenvolvimento do protocolo, dezesseis registros foram pré-classificados manualmente como padrão de referência (*gold standard*) para possibilitar a verificação posterior dos critérios de triagem.

A unidade de análise da auditoria é o instrumento de mensuração, e não o artigo individual. Por essa razão, os 109 artigos foram agregados em clusters de instrumentos com base em um princípio de granularidade controlada. Artigos foram agrupados como usuários do mesmo instrumento quando compartilhavam método de construção, dimensões contempladas,

fonte de dados e estratégia de validação. Por outro lado, instrumentos foram mantidos em clusters distintos quando os próprios autores apresentavam diferenças metodológicas substantivas, como dicionários customizados de transformação digital, abordagens distintas de mineração textual ou diferentes métodos de agregação dos indicadores.

4. Análise dos Resultados e Discussão

4.1 Síntese das revisões: estrutura dimensional e pressupostos

A síntese das onze revisões revelou uma estrutura dimensional relativamente estável para a TD. Apesar das diferenças observadas entre definições, enfoques teóricos e estratégias de categorização, verificou-se convergência quanto aos principais componentes que compõem o fenômeno. Foram identificadas seis dimensões centrais: tecnológica, estratégica, organizacional e estrutural, processos, ecossistema e relações, e pessoas e competências. Além dessas, três dimensões apareceram de forma menos recorrente nas revisões analisadas: efeitos distributivos sobre a força de trabalho, sustentabilidade e efeitos distributivos e societais.

Essa convergência constitui um dos principais resultados da síntese, pois indica que a literatura apresenta maior estabilidade em relação às dimensões da TD do que em relação às suas definições. Essa estrutura constituiu a principal base para a derivação do critério C3 (cobertura dimensional), além de subsidiar aspectos dos critérios C2 (conteúdo) e C5 (registro multinível).

As dimensões variam de forma independente entre as firmas. Uma organização pode avançar tecnologicamente sem se reorganizar internamente, reestruturar sua governança sem adquirir novas capacidades tecnológicas, redefinir seus critérios de criação de valor mantendo processos operacionais tradicionais ou desenvolver competências digitais em sua força de trabalho permanecendo estrategicamente inerte. A variação independente entre indicadores que, em conjunto, constituem o construto é a característica definidora de uma especificação formativa (Boyd et al., 2013; Diamantopoulos & Winklhofer, 2001; Edwards & Bagozzi, 2000; Jarvis et al., 2003). Tratar a TD como um construto reflexivo elimina a estrutura dimensional que a síntese identificou como constitutiva do fenômeno e a substitui por uma suposição de covariação que, quando examinada empiricamente, é frequentemente contrariada pelos dados.

Além da convergência dimensional, a síntese revelou três pressupostos recorrentes associados à forma como a TD é conceituada e mensurada na literatura. Esses pressupostos não representam críticas externas ao campo, mas diagnósticos formulados pelas próprias revisões analisadas.

O primeiro pressuposto refere-se à monotonicidade, isto é, à ideia implícita de que níveis mais elevados de TD produzem necessariamente resultados superiores. Embora diversas revisões reconheçam efeitos negativos, tensões, paradoxos e consequências não intencionais da digitalização, a maior parte dos instrumentos continua operando com uma lógica de progressão unidirecional, na qual mais TD é interpretada como melhor TD. Essa constatação contribuiu para a formulação do critério C4.

O segundo pressuposto consiste no tratamento da firma como unidade suficiente de análise. A maior parte das revisões concentra-se nos efeitos organizacionais da TD, enquanto aspectos relacionados a indivíduos, ecossistemas, relações interorganizacionais e impactos sociais permanecem periféricos. Mesmo quando esses elementos são reconhecidos, tendem a ser tratados como extensões do fenômeno, e não como componentes integrantes da sua mensuração. Esse diagnóstico subsidiou a formulação do critério C5.

O terceiro pressuposto corresponde à interpretação da não adoção digital como falha, resistência ou déficit de implementação. Casos como Kodak e Blockbuster são frequentemente utilizados para ilustrar os riscos da inércia organizacional, e barreiras à TD são normalmente

descritas como obstáculos a serem superados. Entretanto, nenhuma das revisões analisadas distingue explicitamente a não adoção estratégica, entendida como uma decisão deliberada de posicionamento competitivo, da incapacidade de implementar uma estratégia digital desejada. Como resultado, os dois estados permanecem operacionalmente fundidos nos instrumentos predominantes da literatura. Esse diagnóstico fundamentou diretamente o critério C4.

Em conjunto, esses pressupostos evidenciam que parte dos problemas observados na mensuração da TD decorre das técnicas utilizadas e das concepções implícitas incorporadas aos próprios instrumentos. A tradução desses diagnósticos em critérios de auditoria constitui uma das principais contribuições da síntese realizada neste estudo. Ao tornar explícitos pressupostos que frequentemente permanecem implícitos nas operacionalizações existentes, os critérios C1–C5 ampliam a capacidade de avaliar a adequação dos instrumentos empregados na literatura e sua coerência com as definições teóricas que sustentam o construto de TD.

Os resultados da síntese foram traduzidos em cinco critérios de adequação da mensuração, utilizados posteriormente como dimensões de auditoria dos artigos empíricos. Esses critérios estabelecem a ligação entre os achados teóricos da Fase 1 e a avaliação das operacionalizações observadas na literatura. Enquanto as famílias definicionais deram origem aos critérios C1 e C2, a estrutura dimensional consolidada fundamentou o critério C3. Os pressupostos recorrentes identificados nas revisões subsidiaram os critérios C4 e C5. A operacionalização desses critérios como dimensões de auditoria é apresentada na Tabela 2, e sua aplicação aos instrumentos identificados no corpus é reportada nas subseções seguintes.

4.2 Auditoria dos instrumentos empíricos

O processo de triagem da Fase 2 resultou em 109 estudos primários que operacionalizam a TD no nível da firma. A análise dos instrumentos utilizados nesses estudos, considerando seu método de construção, as dimensões contempladas, a fonte de dados empregada e a estratégia de validação adotada, permitiu identificar 14 famílias distintas de instrumentos organizadas em três camadas funcionais. A camada A reúne os 90 estudos que aplicam medidas em pesquisas empíricas (econometria em painel, surveys com análise multivariada, MCDA aplicado a múltiplas firmas ou estudos de caso); a camada B agrega os 12 estudos que propõem modelos de maturidade digital (DMMs) sem aplicação empírica relevante (oferta de instrumentos); e a camada C contém os 7 estudos de metacorpus que analisam como a TD é mensurada (metaevidência sobre o próprio campo). A organização das subseções seguintes segue uma classificação por famílias de instrumentos (text mining, proxies de bases de dados, índices compostos e escalas validadas, DMMs aplicados, propostas conceituais e metacorpus), pois os artigos pertencentes a cada família compartilham mais do que o método empregado: compartilham pressupostos implícitos sobre o que constitui a TD.

O text mining genérico (I-01, 24 artigos) não informa o dicionário utilizado nem apresenta procedimentos de validação; as variantes baseadas em aprendizado de máquina (I-02, 7 artigos) e ponderação por entropia (I-03, 4 artigos) aumentam a sofisticação procedimental sem alterar o pressuposto teórico subjacente; as variantes de processamento de linguagem natural (PLN) com dicionários customizados (I-04, 3 artigos) tratam de forma mais explícita a escolha do vocabulário, mas deixam em aberto a questão do que efetivamente está sendo mensurado.

Nos quatro subclusters, o pressuposto teórico é o mesmo: a visibilidade da temática digital nas comunicações oficiais da firma constitui um proxy válido da TD em curso. Esse pressuposto está implícito na escolha da fonte de dados e não é explicitamente defendido em nenhum dos 38 artigos. A concentração empírica também é estreita: praticamente todos os estudos dessa família utilizam empresas chinesas listadas em ações A, analisadas por meio de

dados em painel, frequentemente com períodos de observação e variáveis dependentes sobrepostos. Quando procedimentos de validação são reportados, estes são predominantemente procedimentais (sensibilidade a especificações econométricas alternativas, IV, GMM e PSM), e não substantivos (validade de construto ou independência dimensional).

Dez estudos do corpus não constroem instrumentos próprios, utilizando índices previamente desenvolvidos por instituições externas, como o *Peking University DT Index of Commercial Banks* (I-05, 7 artigos), índices disponibilizados por bases comerciais, como a CSMAR (I-06, 2 artigos), e índices produzidos por órgãos governamentais nacionais (I-07, 1 artigo). Nesses casos, a validade do instrumento é assumida a partir da instituição responsável pela sua construção, sendo o índice tratado como uma variável pronta para utilização empírica. O subcluster I-05 destaca-se por sua recorrência na literatura sobre TD bancária na China, apresentando uma estrutura uniforme composta por três dimensões (cognição estratégica, organização e produtos).

Vinte e três estudos utilizam instrumentos multidimensionais construídos na forma de índices compostos desenvolvidos pelos próprios autores ou de escalas psicometricamente validadas. Os índices compostos autorais (I-08, 12 artigos) costumam ser acompanhados por testes de robustez econométrica, mas apresentam estratégias heterogêneas de validação do construto; além disso, cada índice é desenvolvido para responder a uma questão empírica específica, contribuindo para a fragmentação metodológica observada no campo. As escalas psicometricamente validadas (I-09, 9 artigos) seguem protocolos formais, incluindo análise fatorial exploratória, alfa de Cronbach, validade de conteúdo por meio do coeficiente V de Aiken, método Delphi e modelagem de equações estruturais. Esse grupo representa o subconjunto metodologicamente mais estruturado do corpus, embora a maioria das escalas tenha sido desenvolvida para contextos setoriais específicos, limitando a comparabilidade entre amostras e estudos. A reutilização de escalas previamente existentes (I-10, 2 artigos) é pouco frequente, indicando que a literatura produz novos instrumentos em ritmo superior ao de sua adoção em pesquisas subsequentes.

Dezessete estudos aplicam modelos de maturidade digital (*Digital Maturity Models* – DMMs), derivados da tradição dos modelos de maturidade de capacidades e dos modelos evolucionários de estágios de desenvolvimento. Os DMMs aplicados a uma ou poucas organizações (I-11, 10 artigos) combinam instrumentos setoriais com técnicas como Processo Analítico Hierárquico (Analytic Hierarchy Process – AHP), avaliação fuzzy ou Análise Multicritério para Apoio à Decisão (*Multi-Criteria Decision Analysis* – MCDA) para atribuição de escores. Nesses casos, a adequação do instrumento é geralmente demonstrada pela aderência ao contexto analisado, enquanto a comparabilidade entre organizações permanece limitada em razão da especificidade de cada modelo. Os DMMs baseados em MCDA aplicados a múltiplas organizações (I-12, 7 artigos) produzem rankings comparativos por meio de técnicas multicritério, mas permanecem fortemente vinculados a contextos setoriais específicos, com validação predominantemente baseada em análises de sensibilidade dos resultados. Em ambos os subclusters, o pressuposto subjacente é o de maturidade de processos, segundo o qual a TD corresponde a uma trajetória evolutiva composta por estágios sucessivos. Essa concepção difere da estrutura multidimensional formativa identificada pela síntese da Fase 1 como constitutiva do construto.

4.3 Afinal, medimos transformação digital?

Os resultados da auditoria indicam que a literatura empírica ainda não mensura a TD de forma consistente com a definição teórica predominante identificada na síntese das revisões. Entre os 90 artigos empíricos da Camada A do corpus auditado, nenhum dos instrumentos

identificados satisfaz simultaneamente os cinco critérios de adequação da mensuração derivados da Fase 1. A Tabela 3 sintetiza os padrões observados nas diferentes famílias de instrumentos.

Tabela 3.

Adequação das famílias de instrumentos aos critérios C1–C5.

Família de instrumentos	n	C1 compromisso com a definição	C2 gatilho vs. conteúdo	C3 cobertura multidimensional formativa	C4 não adoção estratégica	C5 registro multinível
Text mining (I-01 a I-04)	38	Não declarado	Captura o gatilho	Unidimensional ascendente	Confundido	Apenas firma
Proxies de bases de dados (I-05 a I-07)	10	Não declarado	Captura o gatilho	Unidimensional ascendente	Confundido	Apenas firma
Índices compostos (I-08)	12	Implícito	Misto (gatilho e conteúdo)	Multidimensional, validação limitada	Confundido	Apenas firma
Escalas psicométricas (I-09 e I-10)	11	Implícito	Captura conteúdo	Multidimensional, reflexiva	Confundido	Apenas firma
DMMs aplicados (I-11 e I-12)	17	Implícito	Captura conteúdo	Multidimensional, baseada em estágios	Confundido	Apenas firma
DMMs conceituais (I-13)	12	Variável	Variável	Multidimensional, não validada	Não tratado	Apenas firma
Metacorpus (I-14)	7	Revisões sobre mensuração; não constituem instrumentos de mensuração direta				

O principal resultado da auditoria é que diferentes famílias de instrumentos apresentam limitações distintas e recorrentes frente aos critérios derivados da síntese das revisões. Nenhuma delas atende simultaneamente aos cinco critérios de adequação da mensuração, indicando espaço para o desenvolvimento de instrumentos mais aderentes à definição teórica do construto e aos padrões de validade discutidos na literatura metodológica (Boyd, Bergh, Ireland, & Ketchen, 2013; Suddaby, 2010; Venkatraman & Grant, 1986).

Nos instrumentos baseados em text mining e proxies de bases de dados (48 artigos), observa-se a substituição do conteúdo da TD por seus gatilhos tecnológicos. Esses estudos utilizam frequência de termos digitais em relatórios corporativos, investimentos em TI ou indicadores de adoção tecnológica para construir medidas agregadas. O resultado é uma medida da visibilidade da atividade digital na firma, seja em termos discursivos ou financeiros, e não da mudança organizacional que a literatura define como o núcleo do construto (Hanelt et al., 2021). Já os instrumentos multidimensionais, como índices compostos, escalas psicométricas e DMMs aplicados (40 artigos), capturam aspectos mais próximos do conteúdo da TD, mas o fazem por meio de especificações reflexivas ou baseadas em estágios sucessivos. Nesses casos, o que se obtém é uma medida da intensidade percebida das iniciativas digitais ou do grau de maturidade organizacional, e não uma representação da estrutura multidimensional formativa identificada na síntese das revisões. Em ambos os casos, o que os instrumentos capturam é uma manifestação, um antecedente ou uma consequência da TD, e não a mudança organizacional multidimensional que a definição teórica predominante coloca no núcleo do construto.

Além dessas diferenças entre famílias de instrumentos, existe uma limitação comum a todo o corpus. Os instrumentos registram predominantemente efeitos observáveis no nível da firma, porque é nesse nível que os dados estão disponíveis. Como consequência, dimensões relacionadas à reconfiguração de ecossistemas, aos efeitos distributivos sobre o trabalho, às relações com fornecedores e parceiros, aos impactos societais e à sustentabilidade permanecem pouco representadas ou ausentes das medidas utilizadas (Hausberg et al., 2019; Jedynak et al., 2021; Reis & Melão, 2023).

5. Conclusão

Se a pergunta que orientou este estudo era saber se o campo mede aquilo que afirma medir, sua resposta desloca o eixo do debate. Durante boa parte da última década, a agenda de pesquisa em TD concentrou-se em estimar efeitos sobre desempenho, sobre modelos de negócio, sobre vantagem competitiva. O que este estudo sugere é que essa agenda pode ter avançado sobre um alicerce instável: quando o instrumento não corresponde ao construto, os efeitos estimados falam menos sobre o fenômeno e mais sobre aquilo que o instrumento é capaz de enxergar. A contribuição central do trabalho é, nesse sentido, deslocar a conversa do campo do plano dos resultados para o plano das condições que os tornam interpretáveis.

Desse deslocamento derivam as contribuições do estudo. Ao submeter a mensuração da TD ao mesmo escrutínio de validade de construto que a metodologia da gestão estratégica cobra de outros construtos há quase quatro décadas, o trabalho oferece ao campo não apenas um diagnóstico, mas um vocabulário para conversar sobre suas próprias medidas: os cinco critérios de adequação (C1 a C5) funcionam como um protocolo compartilhável, que permite a pesquisadores, pareceristas e editores interrogar um instrumento antes de aceitá-lo como válido. Trata-se de uma contribuição menos sobre o que a TD é e mais sobre como o campo pode saber se está, de fato, falando dela.

A consequência mais desconfortável dessa leitura é retrospectiva. Se parte expressiva do que se acumulou sobre TD e desempenho foi construída sobre medidas de visibilidade ou de maturidade, então uma fração do conhecimento tido como consolidado talvez descreva os instrumentos, e não o fenômeno. Isso não invalida a literatura existente, mas convida a relê-la com outra pergunta em mente: o que exatamente estava sendo medido quando se afirmou medir TD? Reabrir essa pergunta não é um retrocesso, é a condição para que o conhecimento do campo se torne cumulativo em vez de apenas volumoso.

5.1 Limitações

Os resultados devem ser lidos à luz de algumas limitações. A síntese da Fase 1 apoiou-se em onze revisões de referência que, embora abranjam as principais tradições do campo, não esgotam a literatura de revisão, de modo que revisões setoriais, regionais ou em outros idiomas podem incorporar ênfases não contempladas. A auditoria da Fase 2 analisou as práticas de mensuração dos estudos, e não a replicação de suas estimativas originais; assim, as conclusões sobre o alinhamento entre construto e medida não se estendem automaticamente à validade dos resultados substantivos de cada estudo. Por fim, os cinco critérios de adequação, ainda que derivados de forma sistemática, incorporam escolhas interpretativas sobre o que constitui uma mensuração adequada, escolhas que outros pesquisadores poderiam calibrar de maneira distinta.

5.2 Agenda de pesquisas futuras

A principal implicação da auditoria para a pesquisa futura é construtiva: os cinco critérios de adequação não apenas diagnosticam as limitações dos instrumentos existentes, mas oferecem especificações de projeto para o desenvolvimento de medidas mais aderentes à definição teórica do construto. Uma agenda de mensuração orientada por esses critérios abre várias frentes de investigação.

A primeira frente diz respeito ao desenvolvimento de instrumentos formativos que tratem a TD como um construto multidimensional cujas dimensões (tecnológica, estratégica, organizacional, de processos, de ecossistema e de pessoas) sejam constitutivas e não substituíveis, superando as especificações reflexivas que comprometem a independência entre dimensões (C2 e C3). Um instrumento assim concebido mediria o conteúdo da TDe não seus gatilhos tecnológicos, e preservaria a variação independente entre dimensões que a estrutura formativa requer.

A segunda frente refere-se à separação, no próprio desenho da medida, entre a não adoção estratégica e a falha de implementação (C4), de modo que a ausência de determinada iniciativa digital possa ser interpretada como escolha de posicionamento, e não como deficiência de execução. Essa distinção, sistematicamente ausente dos instrumentos auditados, é condição para que a medida discrimine trajetórias organizacionais qualitativamente distintas.

A terceira frente envolve a ampliação do nível de registro para além da firma (C5), incorporando as dimensões interorganizacionais (reconfiguração de ecossistemas, relações com fornecedores e parceiros, efeitos distributivos sobre o trabalho e impactos societais e de sustentabilidade) que permanecem ausentes ou sub-representadas nos instrumentos atuais justamente porque os dados estão mais disponíveis no nível da firma. Superar essa limitação exige o desenvolvimento de fontes e estratégias de coleta que capturem a natureza sistêmica do fenômeno.

Uma via promissora para operacionalizar esses requisitos, e que se coloca como desdobramento direto deste estudo, é a construção de um índice formativo de intensidade da TD, extraído do conteúdo de divulgações corporativas padronizadas com apoio de técnicas de processamento de linguagem natural. Uma abordagem dessa natureza permitiria conciliar a validade de conteúdo, hoje restrita a escalas validadas de pequena amostra, com a escalabilidade requerida por estudos longitudinais e multissetoriais, e constitui uma agenda de desenvolvimento e validação a ser perseguida em trabalhos futuros. De modo mais amplo, o protocolo de auditoria e o conjunto de critérios aqui propostos são transferíveis a outros construtos multidimensionais da gestão estratégica que enfrentam o mesmo dilema entre validade de conteúdo e viabilidade empírica, oferecendo um caminho para fortalecer as bases de mensuração do campo.

Em síntese, o avanço da pesquisa em TD depende menos da produção de novas evidências e mais da garantia de que essas evidências estejam efetivamente alinhadas ao fenômeno que a literatura busca compreender. Enquanto a validade de construto permanecer uma aspiração editorial, e não uma prática rotineira incorporada ao desenho das medidas e exigida de forma consistente pela avaliação por pares, o conhecimento acumulado sobre TD e desempenho continuará vulnerável à distância entre o que se pretende medir e o que efetivamente se mensura.

6. Referências

Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: Toward a next generation of insights. MIS Quarterly, 37(2), 471–482.

Boyd, B. K., Bergh, D. D., Ireland, R. D., & Ketchen, D. J., Jr. (2013). Constructs in strategic management. Organizational Research Methods, 16(1), 3–14. https://doi.org/10.1177/1094428112471298

Boyd, B. K., Gove, S., & Hitt, M. A. (2005). Construct measurement in strategic management research: Illusion or reality? Strategic Management Journal, 26(3), 239–257. https://doi.org/10.1002/smj.444

Crook, T. R., Shook, C. L., Morris, M. L., & Madden, T. M. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. Organizational Research Methods, 13(1), 192–206. https://doi.org/10.1177/1094428109334368

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. Journal of Marketing Research, 38(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Edwards, J. R. (2001). Multidimensional constructs in organizational behavior research: An integrative analytical framework. Organizational Research Methods, 4(2), 144–192. https://doi.org/10.1177/109442810142004

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. Psychological Methods, 5(2), 155–174. https://doi.org/10.1037/1082-989X.5.2.155

Feliciano-Cestero, M. M., Ameen, N., Kotabe, M., Paul, J., & Signoret, M. (2023). Is digital transformation threatened? A systematic literature review of the factors influencing firms' digital transformation and internationalization. Journal of Business Research, 157, 113546. https://doi.org/10.1016/j.jbusres.2022.113546

Godfrey, P. C., & Hill, C. W. L. (1995). The problem of unobservables in strategic management research. Strategic Management Journal, 16(7), 519–533. https://doi.org/10.1002/smj.4250160703

Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2021). A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change. Journal of Management Studies, 58(5), 1159–1197. https://doi.org/10.1111/joms.12639

Hausberg, J. P., Liere-Netheler, K., Packmohr, S., Pakura, S., & Vogelsang, K. (2019). Research streams on digital transformation from a holistic business perspective: A systematic literature review and citation network analysis. Journal of Business Economics, 89(8–9), 931–963. https://doi.org/10.1007/s11573-019-00956-z

Hess, T., Matt, C., Benlian, A., & Wester, F. (2016). Options for formulating a digital transformation strategy. MIS Quarterly Executive, 15(2), 123–139.

Hitt, M. A., Boyd, B. K., & Li, D. (2004). The state of strategic management research and a vision of the future. In D. J. Ketchen & D. D. Bergh (Eds.), Research methodology in strategy and management (Vol. 1, pp. 1–31). Elsevier.

Howell, R. D., Breivik, E., & Wilcox, J. B. (2007). Reconsidering formative measurement. Psychological Methods, 12(2), 205–218. https://doi.org/10.1037/1082-989X.12.2.205

Jamaludin, S. N., & Basari, A. S. H. (2022). Digital transformation strategy: A systematic literature review. Social and Management Research Journal, 19(1), 35–56. https://doi.org/10.24191/smrj.v19i1.17243

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. Journal of Consumer Research, 30(2), 199–218. https://doi.org/10.1086/376806

Jedynak, M., Czakon, W., Kuzniarska, A., & Mania, K. (2021). Digital transformation of organizations: What do we know and where to go next? Journal of Organizational Change Management, 34(3), 629–652. https://doi.org/10.1108/JOCM-10-2020-0336

Ketchen, D. J., Jr., Boyd, B. K., & Bergh, D. D. (2008). Research methods in strategic management: Past accomplishments and future challenges. Organizational Research Methods, 11(4), 643–658. https://doi.org/10.1177/1094428108319843

Ketchen, D. J., Jr., Ireland, R. D., & Baker, L. T. (2013). The use of archival proxies in strategic management studies: Castles made of sand? Organizational Research Methods, 16(1), 32–42. https://doi.org/10.1177/1094428112459911

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. MIS Quarterly, 35(2), 293–334. https://doi.org/10.2307/23044045

Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. Business & Information Systems Engineering, 57(5), 339–343. https://doi.org/10.1007/s12599-015-0401-5

Nadkarni, S., & Prügl, R. (2021). Digital transformation: A review, synthesis and opportunities for future research. Management Review Quarterly, 71(2), 233–341. https://doi.org/10.1007/s11301-020-00185-7

Pfeffer, J. (1993). Barriers to the advance of organizational science: Paradigm development as a dependent variable. Academy of Management Review, 18(4), 599–620. https://doi.org/10.5465/amr.1993.9402210152

Podsakoff, N. P., Shen, W., & Podsakoff, P. M. (2006). The role of formative measurement models in strategic management research: Review, critique, and implications for future research. In D. J. Ketchen & D. D. Bergh (Eds.), Research methodology in strategy and management (Vol. 3, pp. 197–252). Elsevier.

Reis, J., & Melão, N. (2023). Digital transformation: A meta-review and guidelines for future research. Heliyon, 9(1), e12834. https://doi.org/10.1016/j.heliyon.2023.e12834

Reis, J., Amorim, M., Melão, N., & Matos, P. (2018). Digital transformation: A literature review and guidelines for future research. In Á. Rocha, H. Adeli, L. Reis, & S. Costanzo (Eds.), Trends and advances in information systems and technologies (pp. 411–421). Springer. https://doi.org/10.1007/978-3-319-77703-0_41

Schallmo, D., Williams, C. A., & Boardman, L. (2017). Digital transformation of business models: Best practice, enablers, and roadmap. International Journal of Innovation Management, 21(8), 1740014. https://doi.org/10.1142/S136391961740014X

Sebastian, I. M., Ross, J. W., Beath, C., Mocker, M., Moloney, K. G., & Fonstad, N. O. (2017). How big old companies navigate digital transformation. MIS Quarterly Executive, 16(3), 197–213.

Suddaby, R. (2010). Editor's comments: Construct clarity in theories of management and organization. Academy of Management Review, 35(3), 346–357. https://doi.org/10.5465/amr.35.3.zok346

Venkatraman, N., & Grant, J. H. (1986). Construct measurement in organizational strategy research: A critique and proposal. Academy of Management Review, 11(1), 71–87. https://doi.org/10.5465/amr.1986.4282628

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. Journal of Business Research, 122, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. Journal of Strategic Information Systems, 28(2), 118–144. https://doi.org/10.1016/j.jsis.2