Supplementary File S2. Search Strategy, Screening Criteria, and Coding Protocol

This supplementary file provides the detailed search strategy, screening criteria, coding protocol, effect-size extraction rules, and synthesis procedures used in the systematic review-based meta-analysis. It should be read together with Supplementary File 1, which provides the PRISMA 2020 checklist, and Supplementary File 4, which provides the PRISMA-S checklist.

1. Search Sources and Search Window

The literature search was conducted in Web of Science (Clarivate), Scopus (Elsevier), and ProQuest ABI/INFORM on the ProQuest platform. These databases were selected because they provide broad coverage of peer-reviewed journal articles in business, management, operations, decision sciences, economics, information systems, social sciences, and related fields. The search window covered studies published from 2000 to June 2026. The initial search was conducted in May 2025 and was subsequently updated in June 2026 to incorporate recent developments in digital transformation, emerging digital technologies, and supply chain resilience research.

The search was restricted to peer-reviewed English-language journal articles. Search terms were applied to the metadata fields available in each database, including title, abstract, keywords, topic, subject, or summary fields depending on database-specific syntax. The search strategy was organized around two evidence pools. The DT–SCR evidence pool was identified by combining digital transformation-related keywords with supply chain resilience-related keywords. The SCR–performance evidence pool was identified by combining supply chain resilience-related keywords with firm performance-related keywords.

Search results from the two evidence pools were consolidated before duplicate removal and screening. The consolidated database search identified 4,479 records before deduplication, as reported in Figure 2 of the main manuscript. No study registries, organizational websites, conference proceedings, or author-contact procedures were used as formal information sources.

2. Keyword Families

2.1 Digital transformation-related keywords

The digital transformation keyword family included terms referring to digital transformation, digitalization, digital technologies, digital capabilities, and major digital technology applications in supply chain management and operations. Specifically, this keyword family included: digital transformation, digitalization, digital technologies, digital technology, digital capabilities, digital capability, Industry 4.0, Industry 5.0, artificial intelligence, generative artificial intelligence, machine learning, Internet of Things, blockchain, cloud computing, big data analytics, digital platforms, digital twins, cyber-physical systems, and smart manufacturing.

2.2 Supply chain resilience-related keywords

The supply chain resilience keyword family captured resilience-related concepts associated with disruption management, recovery, adaptation, and continuity in supply chain contexts. The terms used in this keyword family were: supply chain resilience, supply chain disruption, supply chain risk, supply chain robustness, supply chain recovery, supply chain adaptability, supply chain agility, and supply chain continuity.

2.3 Performance-related keywords

The performance keyword family captured firm-level, operational, financial, market, and supply-chain-related performance outcomes. The terms used in this keyword family were firm performance,

operational performance, financial performance, market performance, competitive advantage, supply chain performance, and economic performance.

3. Database-Specific Search Strategy

The same keyword logic was applied across Web of Science, Scopus, and ProQuest ABI/INFORM, with syntax adapted to the search fields and operators used by each database. Two searches were conducted in each database: one for the DT–SCR evidence pool and one for the SCR–performance evidence pool.

3.1 Web of Science

In Web of Science, searches were conducted in the Topic field. The DT–SCR evidence pool was searched using the following string:

TS = (("digital transformation" OR digitalization OR "digital technologies" OR "digital technology" OR "digital capabilities" OR "digital capability" OR "Industry 4.0" OR "Industry 5.0" OR "artificial intelligence" OR "generative artificial intelligence" OR "machine learning" OR "Internet of Things" OR blockchain OR "cloud computing" OR "big data analytics" OR "digital platform*" OR "digital twin*" OR "cyber-physical systems" OR "smart manufacturing") AND ("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability" OR "supply chain agility" OR "supply chain continuity"))

The SCR–performance evidence pool was searched using the following string:

TS = (("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability" OR "supply chain agility" OR "supply chain continuity") AND ("firm performance" OR "operational performance" OR "financial performance" OR "market performance" OR "competitive advantage" OR "supply chain performance" OR "business performance" OR "economic performance"))

The Web of Science search was limited to English-language journal articles published between 2000 and June 2026. The final updated search was conducted in June 2026.

3.2 Scopus

In Scopus, searches were conducted in the title, abstract, and keyword fields. The DT–SCR evidence pool was searched using the following string:

TITLE-ABS-KEY ((("digital transformation" OR digitalization OR "digital technologies" OR "digital technology" OR "digital capabilities" OR "digital capability" OR "Industry 4.0" OR "Industry 5.0" OR "artificial intelligence" OR "generative artificial intelligence" OR "machine learning" OR "Internet of Things" OR blockchain OR "cloud computing" OR "big data analytics" OR "digital platform*" OR "digital twin*" OR "cyber-physical systems" OR "smart manufacturing") AND ("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability" OR "supply chain agility" OR "supply chain continuity")))

The SCR–performance evidence pool was searched using the following string:

TITLE-ABS-KEY (("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability") AND ("firm performance" OR "operational performance" OR "financial performance" OR "market performance" OR "competitive advantage" OR "supply chain performance" OR "economic performance"))

The Scopus search was limited to English-language journal articles published between 2000 and June 2026. The final updated search was conducted in June 2026.

3.3 ProQuest ABI/INFORM

In ProQuest ABI/INFORM, searches were conducted using abstract, title, or summary fields depending on database availability. The DT–SCR evidence pool was searched using the following string:

(("digital transformation" OR digitalization OR "digital technologies" OR "digital capabilities" OR "Industry 4.0" OR "Industry 5.0" OR artificial intelligence OR "generative artificial intelligence" OR "Internet of Things" OR blockchain OR "cloud computing" OR "big data analytics" OR "digital platform*" OR "digital twin*" OR "cyber-physical systems") AND ("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability"))

The SCR–performance evidence pool was searched using the following string:

(("supply chain resilience" OR "supply chain disruption" OR "supply chain risk" OR "supply chain robustness" OR "supply chain recovery" OR "supply chain adaptability") AND ("firm performance" OR "operational performance" OR "financial performance" OR "market performance" OR "competitive advantage" OR "supply chain performance" OR "economic performance"))

The ProQuest ABI/INFORM search was limited to peer-reviewed English-language journal articles published between 2000 and June 2026. The final updated search was conducted in June 2026.

4. Search Consolidation, Deduplication, and Study Selection

The DT–SCR and SCR–performance searches were conducted separately in Web of Science, Scopus, and ProQuest ABI/INFORM. The search results from both evidence pools were subsequently combined into a unified screening process.

Because some articles appeared in multiple databases and some studies contributed evidence to both the DT–SCR and SCR–performance evidence pools, pool-specific counts were not directly summed to represent the number of unique studies. Duplicate records were identified and removed using EndNote and Excel, followed by manual verification based on bibliographic information, including title, authors, publication year, journal, and DOI where available.

The consolidated search identified 4,479 records before screening. After duplicate removal, 2,187 records remained for title and abstract screening. Following title and abstract screening, 423 full-text articles were assessed for eligibility. After full-text assessment, 223 articles were excluded according to the predefined eligibility criteria.

The final sample included 200 unique studies. Among these studies, 124 studies contributed to the DT–SCR evidence pool, 104 studies contributed to the SCR–performance evidence pool, and 28 studies contributed evidence to both pools.

5. Eligibility Criteria and Screening Protocol

The screening process followed the PRISMA 2020 framework (Page et al., 2021), consisting of record identification, duplicate removal, title and abstract screening, full-text eligibility assessment, and final inclusion. The same eligibility criteria were applied to both evidence pools.

To improve consistency in study selection and reduce subjective judgments, predefined decision rules were established before full-text screening. During eligibility assessment, studies were evaluated sequentially based on whether they represented empirical research, whether digital transformation and resilience were examined as formal research constructs, whether an empirical relationship between digital transformation and resilience was examined, and whether sufficient statistical information was available for meta-analysis.

Inclusion criteria. Studies were retained when they met all of the following criteria:

(1) they were peer-reviewed English-language empirical journal articles;

(2) they were conducted in firm, organizational, supply chain, or interorganizational supply chain contexts;

(3) they examined either the DT–SCR relationship, the SCR–performance relationship, or both;

(4) they reported sufficient statistical information to compute or convert an effect size; and

(5) they provided sample size information.

For the DT–SCR evidence pool, digital transformation and resilience were required to represent formal constructs rather than background concepts or technical tools. Digital transformation was considered a formal construct when it was modeled as a variable, capability, resource, antecedent, mediator, moderator, configuration condition, or formally measured construct. Resilience was considered a formal construct when it was modeled as a variable, outcome, mediator, moderator, or formally measured capability.

When studies did not specify a formal DT→SCR structural path but reported an extractable zero-order association between formally measured digital transformation and resilience constructs, they were retained as correlational evidence if the sample size was identifiable and the association was compatible with the meta-analytic framework. Potentially overlapping samples were identified during full-text screening. Studies with suspected shared datasets were flagged for further verification rather than being immediately treated as independent evidence.

Exclusion criteria. Studies were excluded when they were reviews, conceptual papers, editorials, book chapters, dissertations, or conference papers. Studies were also excluded if they were not conducted in relevant firm or supply chain contexts; used unclear or incompatible construct definitions; lacked sufficient statistical information for effect-size extraction or conversion; duplicated previously reported datasets without providing independent evidence; or did not provide usable evidence for either focal relationship.

6. Coding Protocol

A predefined coding protocol and decision rules were developed to guide construct identification, relationship classification, and data extraction. Coding covered bibliographic descriptors, sample descriptors, construct descriptors, moderator variables, and effect-size information. For each included study, we coded author, publication year, journal, sample size, country or region, industry context, firm size, and firm age.

Construct identification followed explicit coding rules. Digital transformation was coded as a formal construct only when it represented a measured variable, capability, resource, antecedent, mediator, moderator, or configuration condition. Resilience was coded as a formal construct only when it represented a measured variable or modeled construct rather than a general contextual concept, supplier characteristic, or descriptive term.

The relationship between digital transformation and resilience was classified as Direct, Indirect, Moderated, Configurational, Correlational, No, or Unclear. Direct relationships referred to studies explicitly modeling digital transformation as an antecedent of resilience. Indirect relationships referred to studies where digital transformation affected resilience through mediating mechanisms. Moderated relationships referred to studies where digital transformation influenced the strength of another antecedent–resilience relationship. Configurational relationships referred to studies using configuration approaches in which digital transformation was one of the conditions associated with resilience outcomes. Correlational relationships referred to studies where both constructs were formally measured and an extractable zero-order association was reported without an explicit structural path. Studies where digital

transformation and resilience were only separately discussed or where digital transformation influenced outcomes unrelated to resilience were classified as No.

Two researchers independently coded the studies. Disagreements were resolved through discussion and re-examination of the original articles until agreement was reached. Inter-coder agreement before reconciliation was 96.54%.

6.1 Digital transformation coding

Digital transformation was coded into three categories: digital technologies, digital capabilities, and combined.

Digital technologies refer to the adoption, implementation, or utilization of specific digital tools, technological infrastructures, or digital solutions in organizational and supply chain contexts. Examples include artificial intelligence, Internet of Things, blockchain, cloud computing, big data analytics, digital platforms, digital twins, cyber-physical systems, Industry 4.0-related technologies, Industry 5.0-related technologies, and other emerging digital applications.

Digital capabilities refer to organizational capabilities that enable firms to integrate, orchestrate, leverage, and reconfigure digital resources to support business and operational processes. Examples include digital sensing capability, digital integration capability, data-driven decision-making capability, digital innovation capability, technology orchestration capability, and digital coordination capability.

Studies were coded as digital technologies when the measurement primarily focused on the adoption, use, or implementation of specific digital technologies. Studies were coded as digital capabilities when the measurement emphasized organizational capabilities, resource integration, digital transformation capability, or the ability to leverage digital resources. Studies were coded as combined when digital technology adoption and capability development were simultaneously captured or when the measurement did not clearly distinguish between technology-oriented and capability-oriented dimensions.

Emerging technologies were included only when they were examined as formal digital transformation constructs. Technologies mentioned only as background information, future research directions, analytical tools, simulation environments, or data-processing approaches were not coded as formal digital transformation constructs.

6.2 Supply chain resilience coding

Supply chain resilience was coded as focal-firm-level resilience when studies examined a firm's ability to maintain, restore, adapt, or reconfigure supply chain-related activities and interorganizational coordination under disruptions. SCR operationalization was coded as proactive resilience, reactive resilience, or overall resilience.

Proactive resilience refers to preparedness-oriented capabilities developed before disruptions occur, including anticipation, prevention, risk monitoring, vulnerability identification, buffering, and readiness-building activities. Reactive resilience refers to response-oriented capabilities after disruptions occur, including recovery, adaptation, flexibility, restoration, and post-disruption reconfiguration. Overall resilience refers to studies measuring SCR as a composite construct without distinguishing between proactive and reactive dimensions. Studies were not coded as SCR when resilience appeared only as a general contextual concept, descriptive term, supplier attribute, or background discussion without formal measurement or empirical modeling.

6.3 Firm performance coding

Firm performance was coded into four categories: operational performance, financial performance, market performance, and overall performance.

Operational performance refers to efficiency-, process-, and operations-related outcomes, including productivity, delivery performance, cost efficiency, flexibility, quality, and operational improvement. Financial performance refers to accounting-based or financial outcomes, including profitability, return measures, sales growth, and other financial indicators. Market performance refers to market-oriented outcomes, including customer performance, competitive advantage, market position, innovation-related outcomes, and customer-related outcomes. Overall performance refers to composite performance measures that integrate multiple dimensions without distinguishing among operational, financial, and market outcomes.

When studies reported multiple performance dimensions, each dimension was coded separately for dimension-specific analyses. Overall performance measures were retained for the overall SCR–performance synthesis but were not treated as an independent performance-dimension subgroup in moderator comparisons.

6.4 Moderator coding

Moderator coding followed predefined operational rules.

Firm size was classified using a cutoff of 500 employees. This threshold has been used in prior meta-analytic research to distinguish small and medium-sized enterprises (SMEs) from larger firms (Rosenbusch et al., 2011; Sarooghi et al., 2015). Accordingly, studies reporting a mean workforce of fewer than 500 employees were classified as SME samples, whereas studies reporting a mean workforce of 500 employees or more were classified as large-firm samples. This classification represents a literature-based operationalization for subgroup comparison rather than a universal definition of firm size categories.

Firm age was classified using an eight-year cutoff (Brinckmann, J. et al.,2010). Following Rauch and Hatak (2016), we acknowledge that firm-age thresholds may vary across industries and research contexts because organizational maturity is context dependent. However, prior primary studies and meta-analytic research have used an average firm age of eight years to distinguish young from older firms. Accordingly, studies reporting a mean firm age of eight years or less were classified as young-firm samples, whereas studies reporting a mean firm age above eight years were classified as mature-firm samples. This cutoff was adopted as a literature-informed classification criterion for subgroup analysis rather than a strict organizational life-cycle boundary.

Industry type was coded as manufacturing, services, or Mixed/NA based on the primary industry of sampled firms. Studies were classified as manufacturing when the sample primarily consisted of firms engaged in production, processing, or physical goods manufacturing activities. Studies were classified as services when the sample primarily consisted of service-oriented firms, including financial, healthcare, information, retail, consulting, and other non-manufacturing activities. When studies included firms from multiple industries without a clearly dominant category, or when industry information was insufficient to determine the primary classification, they were coded as Mixed/NA and retained only in the overall synthesis. For descriptive purposes, industry subsectors were also coded when sufficient information was reported. Subsector classifications were based on the primary business activities of sampled firms. Because several subsectors contained a limited number of studies, these analyses were treated as exploratory and were not used for hypothesis testing.

Regional development level was coded as developed or developing based on IMF classification, while mixed or unclear samples were coded as Mixed/NA.

National culture was operationalized using Hofstede's four cultural dimensions: power distance (PDI), individualism–collectivism (IDV), uncertainty avoidance (UAI), and long-term orientation (LTO)

(Hofstede et al., 2010). For single-country samples, the corresponding country score was assigned. High and low groups were created using the mean score across coded studies as the cutoff. Multi-country samples without sufficient weighting or operationalization information were coded as NA. Throughout the coding process, NA indicates information not reported, Overall indicates an undifferentiated construct measure, Combined indicates DT measures covering both technology and capability dimensions, and Mixed indicates samples spanning multiple categories.

7. Effect-Size Extraction and Transformation Protocol

Pearson's correlation coefficient (r) was selected as the common effect-size metric because it allows consistent synthesis across studies using different empirical designs and statistical approaches. When studies reported zero-order correlations for the focal relationship, the reported correlations were extracted directly.

When multiple effect sizes were reported from the same study and sample for the same focal relationship, aggregation rules were applied for the main-effect analyses to avoid over-representing individual samples. Specifically, multiple effects referring to the same construct relationship and based on the same sample were averaged before inclusion in the overall synthesis. However, when effect sizes represented theoretically distinct construct operationalizations, such as different DT dimensions, SCR dimensions, or performance dimensions, dimension-specific effects were retained for subgroup and sensitivity analyses.

7.1 Conversion of statistical information

When zero-order correlations were unavailable, reported statistics were converted into approximate correlation coefficients when sufficient information was provided. Standardized regression coefficients (β) were converted into approximate correlations following Peterson and Brown (2005):

$$r = \beta + 0.05\lambda$$

where λ = 1 when β ≥ 0 and λ = 0 when β < 0.

This conversion was applied only when zero-order correlations were unavailable and the reported coefficients represented comparable bivariate associations between the focal constructs.

When studies reported other convertible statistics (e.g., t-values or F-values) together with sufficient information, effect sizes were calculated using established meta-analytic conversion procedures.

7.2 Reliability correction and Fisher transformation

When reliability estimates were available for the focal constructs, correlations were corrected for attenuation using the reported reliability coefficients. When reliability information was unavailable, reliability values were imputed using the sample-size-weighted mean reliability estimated from studies reporting reliability information for the corresponding construct.

After effect-size extraction and transformation, all correlation coefficients were converted into Fisher's z values for statistical pooling:

$$z = 0.5\ln\frac{1+r}{1-r}$$

Pooled effects were subsequently back-transformed from Fisher's z values into Pearson's r values for interpretation.

7.3 Effect-size independence and synthesis rules

The unit of analysis was determined according to the research question and evidence pool.

For the overall DT–SCR and SCR–performance syntheses, effect sizes were aggregated to maintain independence among observations when multiple effects were reported from the same study sample.

For construct-specific analyses, including DT dimensions (digital technologies versus digital capabilities), SCR operationalizations (proactive versus reactive resilience), and performance dimensions (operational, market, and financial performance), dimension-specific effects were retained because they represent theoretically meaningful differences in construct operationalization.

8. Meta-Analytic Model and Moderator Tests

Random-effects models were used as the primary synthesis approach because the included studies differed in theoretical perspectives, sample characteristics, measurement approaches, research contexts, and empirical designs. Therefore, the true effect sizes were expected to vary across studies rather than represent a single common population effect.

All meta-analytic procedures were conducted using Comprehensive Meta-Analysis (CMA) Version 3.0. Effect sizes were pooled using inverse-variance weighting. The pooled effects were interpreted using back-transformed Pearson's r values.

Between-study heterogeneity was assessed using Cochran's Q statistic and the I^2 statistic. Cochran's Q evaluates whether the observed variation among effect sizes exceeds the variation expected from sampling error alone. The I^2 statistic indicates the proportion of total observed variance attributable to between-study heterogeneity.

Moderator analyses were conducted using subgroup comparisons based on the between-group heterogeneity statistic (Qb). Subgroup analyses were performed only when sufficient studies were available within comparison groups. For the DT–SCR evidence pool, subgroup analyses examined whether the strength of the relationship varied across organizational, contextual, cultural, and construct-related conditions. Organizational and contextual moderators included: firm size; firm age; industry type; regional development level. Cultural moderators included Hofstede's four cultural dimensions: power distance (PDI); individualism–collectivism (IDV); uncertainty avoidance (UAI); long-term orientation (LTO). Construct-related moderators included: digital transformation dimensions: digital technologies; digital capabilities; SCR operationalization: proactive resilience; reactive resilience.

For the DT–SCR evidence pool, construct-related subgroup analyses were conducted at the effect level because they represented differences in construct operationalization within studies. Contextual and cultural subgroup analyses were conducted based on study-level characteristics.

For the SCR–performance evidence pool, subgroup analyses focused on whether the association between SCR and firm performance differed across performance dimensions. Performance outcomes were categorized into: operational performance; market performance; financial performance. Overall performance measures were retained for the overall synthesis but were not included as an independent subgroup category when comparing performance dimensions.

For each subgroup analysis, pooled effects were estimated separately for each category. Differences between subgroup effects were evaluated using the Qb statistic. A statistically significant Qb indicates that the strength of the relationship differs across comparison groups.

9. Publication Bias Assessment

Publication bias was assessed separately for the DT–SCR and SCR–performance evidence pools using Egger's regression test and fail-safe N analyses.

Egger's regression test was conducted to examine whether the distribution of effect sizes showed evidence of asymmetry that could indicate potential small-study effects or publication bias. A non-

significant result suggests that the observed effect distribution does not provide substantial evidence of asymmetry.

Fail-safe N analyses were conducted to estimate the number of additional null-effect studies required to reduce the observed meta-analytic effect to a non-significant level. A larger fail-safe N indicates greater robustness of the pooled effect against the potential influence of unpublished or missing studies.

Publication bias assessments were interpreted together with sensitivity analyses rather than as standalone evidence. Because statistical tests of publication bias can have limited power when the number of included studies is small, results were evaluated considering the overall evidence base and robustness of the pooled effects.

10. Sensitivity Analysis

Leave-one-study-out sensitivity analyses were conducted to assess the robustness of the pooled effects. In each analysis, one study was sequentially removed, and the pooled effect size was recalculated using the random-effects model.

For the DT–SCR evidence pool, the pooled correlations remained stable across all iterations, ranging from r=0.594 to r=0.603. For the SCR–performance evidence pool, the pooled correlations ranged from r=0.584 to r=0.597. These results indicate that no single study had a substantial influence on the overall effect estimates. The sensitivity analyses were interpreted together with the main meta-analytic results to evaluate the robustness of the synthesized relationships.

11. Risk-of-Bias and Certainty-of-Evidence Statements

A separate study-level risk-of-bias assessment was not conducted because the present review synthesizes observational empirical studies in management and supply chain research rather than randomized intervention studies. Instead, potential sources of bias were addressed through multiple methodological safeguards throughout the review process.

Specifically, study quality and potential bias concerns were addressed through predefined eligibility criteria, explicit construct identification rules, independent coding by two researchers, reconciliation of coding disagreements, predefined effect-size extraction and transformation rules, synthesis-level publication-bias assessment, and leave-one-study-out sensitivity analyses. These procedures were designed to reduce subjectivity in study selection, construct classification, and statistical synthesis.

A formal certainty-of-evidence assessment, such as GRADE, was not conducted because GRADE was primarily developed for evaluating evidence from intervention studies and clinical or policy contexts. Instead, the robustness and consistency of the synthesized evidence were evaluated through random-effects models, confidence intervals, heterogeneity statistics, subgroup analyses, publication-bias assessments, and sensitivity analyses.

12. Search Updates and Search Peer Review

The initial database search was conducted in May 2025. To incorporate recent developments in digital transformation, emerging digital technologies, and supply chain resilience research, the search was updated in June 2026.

The updated search followed the same database sources, keyword logic, eligibility criteria, and screening procedures as the initial search. The update incorporated recently published studies and expanded digital transformation-related terms, including emerging concepts such as Industry 5.0, digital twins, and generative artificial intelligence.

The search strategy was not externally peer-reviewed using a formal search-review instrument. However, the search process was documented through database-specific search strings, explicit eligibility

criteria, predefined coding rules, and a transparent screening procedure. These materials are provided in the supplementary files to facilitate reproducibility.

13. Supplementary File Linkage

Supplementary File 1 provides the PRISMA 2020 checklist. Supplementary File 2 provides the search strategy, study selection and eligibility criteria, coding protocol, effect-size extraction rules, and synthesis procedures. Supplementary File 3 provides the coding results and included studies list. Supplementary File 4 provides the PRISMA-S checklist for reporting literature searches.

References

Brinckmann, J.; Grichnik, D.; Kapsa, D. Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning-performance relationship in small firms. *Journal of Business Venturing* **2010**, *25*, 24–40. https://doi.org/10.1016/j.jbusvent.2009.12.002

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and Organizations: Software of the Mind* (3rd ed.). New York, NY: McGraw-Hill.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology, 90*(1), 175–181. https://doi.org/10.1037/0021-9010.90.1.175

Rauch, A., & Hatak, I. (2016). A meta-analysis of different HR-enhancing practices and performance of small and medium-sized firms. Journal of Business Venturing, 31(5), 485–504. https://doi.org/10.1016/j.jbusvent.2016.05.005.

Rosenbusch, N., Brinckmann, J., & Bausch, A. (2011). Is innovation always beneficial? A meta-analysis of the relationship between innovation and performance in SMEs. *Journal of Business Venturing, 26*(4), 441–457. https://doi.org/10.1016/j.jbusvent.2009.12.002

Sarooghi, H., Libaers, D., & Burkemper, A. (2015). Examining the relationship between creativity and innovation: A meta-analysis of organizational, cultural, and environmental factors. *Journal of Business Venturing, 30*(5), 714–731. https://doi.org/10.1016/j.jbusvent.2014.12.003