A Conceptual Framework for the Green Balanced Scorecard: An Expert-Consensus-Based Approach

Mohammadreza Ola, Elham Davari Ghezelhesar, Hamid Erfanian Khanzadeh

Faculty of Accounting ،Islamic Azad University of Mashhad, Mashhad, Iran

Abstract

Despite growing scholarly attention to the Green Balanced Scorecard (GBSC), the literature has not yet converged on an integrated set of indicators that simultaneously covers all four perspectives of this framework. Unlike the Sustainability Balanced Scorecard (SBSC), which encompasses economic, social, and environmental dimensions, the study focuses on the environmental dimension of sustainability to extract GBSC indicators and develop its conceptual framework based on expert consensus. The study was conducted in two stages. First, 47 initial indicators were extracted through document analysis and classified into the four GBSC perspectives. The indicators were then evaluated using a modified Fuzzy Delphi Method in two rounds, with 15 experts participating in the first round and 10 in the second. Data analysis was based on Triangular Fuzzy Numbers, and indicators were accepted according to the criteria of threshold ≤ 0.20, consensus ≥ 80% (the ceiling of 0.75N), and defuzzification ≥ 0.50. Of the 47 initial indicators, 29 were accepted, comprising 8 financial indicators, 4 customer indicators, 10 internal-process indicators, and 7 learning and growth indicators. These indicators formed the four-perspective GBSC framework. The indicators were prioritized using the Composite Ranking Index (CRI), and the convergence of expert judgment across the two Delphi rounds was examined. By conceptually distinguishing GBSC from SBSC and employing Fuzzy Delphi, between-round comparison, and CRI ranking, this study presents a set of environmental indicators validated through expert consensus within a four-perspective conceptual framework. The framework identifies and prioritizes the indicators and may provide a basis for future empirical research on environmental performance measurement; the cause-and-effect relationships between the four perspectives require testing in subsequent studies.

Keywords: Green Balanced Scorecard (GBSC); Fuzzy Delphi Method (FDM); Environmental Performance Indicators; Performance Measurement; Conceptual Framework

Uncorrected proof

1. Introduction

In recent decades, the manufacturing sector has become one of the most important sources of pressure on the environment and, in the European Union, accounts for the second-largest share of greenhouse gas emissions after the energy sector (Panagiotopoulou et al., 2022). At the same time, rising consumption of natural resources and waste generation has reinforced the need to design standardized indicators for monitoring organizations' environmental performance (Voukkali et al., 2023). Empirical evidence also indicates that environmental violations may increase firms' financial distress risk, making environmental conduct a strategic and financial issue (Xing et al., 2025).

In response to these developments, Environmental Management Accounting (EMA) has become an important tool for identifying and reporting environmental costs and benefits alongside financial information. Studies have shown that the use of EMA, together with top-management support and environmental knowledge management, contributes to improved organizational environmental performance (Stanescu et al., 2021; Bresciani et al., 2023). In addition, the disclosure of environmental, social, and governance (ESG) information plays an increasingly important role in stakeholder assessments and firm value (Moussa and Elmarzouky, 2023).

The Balanced Scorecard (BSC), introduced by Kaplan and Norton (1992), evaluates organizational performance across four perspectives: financial, customer, internal processes, and learning and growth. Despite the importance of this framework in strategic management, its classical version did not include a separate indicator for measuring environmental consequences. Accordingly, recent studies have sought to integrate sustainability considerations into the BSC structure (Jassem et al., 2022).

In this context, the Sustainability Balanced Scorecard (SBSC) was developed as an approach for the simultaneous integration of the economic, social, and environmental dimensions. The systematic review by Jassem et al. (2022) shows that the literature presents the SBSC in two main forms: SBSC-4, which integrates sustainability indicators into the four traditional perspectives, and SBSC-5, which adds a separate sustainability perspective. In addition, studies such as Zhou et al. (2020), Bendig et al. (2023), and Martín-Gómez et al. (2024) have shown that BSC logic can be used to assess sustainable energy, green manufacturing, and the circular economy.

Despite the expansion of sustainability-oriented approaches, it is necessary to distinguish between the SBSC and the GBSC. The SBSC is a comprehensive framework for integrating the economic, social, and environmental dimensions of sustainability into an organizational performance-management system (Jassem et al., 2022). In contrast, this

Uncorrected proof

study defines the GBSC as an environmental specialization of the BSC; that is, green objectives and indicators are integrated into the same four traditional BSC perspectives without adding a separate perspective to the structure. Accordingly, the GBSC is specifically focused on measuring and managing environmental performance, thereby avoiding conceptual conflation with the broader SBSC framework.

The literature review indicates that GBSC-related research faces three main gaps: (1) Theoretical gap: Although theories related to the Triple Bottom Line, the Natural Resource-Based View, and stakeholder and legitimacy perspectives have been examined separately, their integration within a four-perspective GBSC framework has been less frequently observed (Dzhengiz and Niesten, 2020; Khan et al., 2024). (2) Methodological gap: Many studies have been limited to extracting indicators from secondary data and literature reviews and have made insufficient use of structured expert judgment to validate indicators (Aljamal et al., 2024). (3) Applied gap: Existing studies have mainly focused on individual dimensions such as the green supply chain, green training, or green supplier selection, while comprehensive sets of environmental indicators that integratively cover all four BSC perspectives have rarely been presented (Abuzaid et al., 2024; Li and Zhong, 2024; Sun et al., 2024).

The present study was conducted to extract GBSC indicators and develop a four-perspective conceptual framework based on expert consensus. The two main research questions are:

What are the key GBSC indicators across the financial, customer, internal-process, and learning and growth perspectives?

How are these indicators prioritized under conditions of linguistic uncertainty in expert judgment?

To answer these questions, 47 initial indicators were first extracted through document analysis and classified into the four perspectives. The indicators were then evaluated using a modified Fuzzy Delphi Method in two rounds, with 15 experts participating in the first round and 10 in the second. The study's contributions can be stated at three levels: (1) theoretical: providing a clear conceptual distinction between GBSC and SBSC, focusing on the environmental dimension of sustainability, and mapping five theoretical foundations onto the four perspectives and their indicators; (2) methodological : combining targeted document analysis with Fuzzy Delphi to reduce linguistic uncertainty and enhance the content validity of the indicators; and (3) applied: providing a four-perspective conceptual framework for organizing and evaluating environmental

uncorrected proof

performance indicators ,with the financial indicators mapped to established accounting treatments and disclosure frameworks, that may serve as a basis for future empirical research and field validation. This study constitutes the first stage of a research program, namely the identification and consensus-based prioritization of indicators; the strategic cause-and-effect relationships between the four perspectives were neither modeled nor tested here and are left to subsequent research.

The article is structured as follows: Section 2 presents the theoretical foundations, the research background, and the conceptual framework; Section 3 describes the Fuzzy Delphi methodology; Section 4 reports the findings; Section 5 provides the discussion; and Sections 6 to 8 present the limitations, recommendations for future research, and the overall conclusion.

2. Theoretical foundations

2.1. Basic concepts and theoretical foundations

The present study is grounded in five theoretical foundations: three main foundations and two complementary perspectives. First, the Balanced Scorecard (BSC), introduced by Kaplan and Norton (1992), enables organizations to monitor past performance and the drivers of future performance simultaneously across the financial, customer, internal-process, and learning and growth perspectives. Second, the Triple Bottom Line (TBL) model proposes the simultaneous assessment of economic, social, and environmental dimensions and, in this study, provides the basis for focusing on the environmental dimension and distinguishing the concept of "green" from "sustainable" (Elkington, 1997). Third, the Natural Resource-Based View (NRBV) (Hart, 1995) regards environmental capabilities and resources as sources of sustainable competitive advantage; studies have shown that managers' environmental competencies, organizational absorptive capacity, and green knowledge management can strengthen organizational learning, green innovation, and environmental performance (Dzhengiz and Niesten, 2020; Khan et al., 2024). Alongside these three main foundations, stakeholder theory (Freeman, 1984) and legitimacy theory (Suchman, 1995) complement the analytical framework of the study. These two perspectives emphasize the need for organizations to respond to stakeholder expectations, maintain social legitimacy, and establish credible mechanisms for environmental governance and reporting (Andreoli et al., 2025; Pajuelo-Moreno et al., 2024). The correspondence between these five foundations, the four GBSC perspectives and the resulting indicators is summarized in Table 1, in Section 2.7.1.

Uncorrected proof

2.2. The balanced scorecard and its evolution into a green version

The classical BSC did not include a separate indicator for measuring environmental consequences; consequently, sustainability-oriented approaches were developed. The systematic review by Jassem et al. (2022) shows that the SBSC has been presented in two principal forms: SBSC-4, which integrates sustainability indicators into the four traditional perspectives, and SBSC-5, which adds a separate sustainability perspective. Recent studies have also extended the application of the BSC to sustainable energy, green manufacturing, and the circular economy (Bendig et al., 2023; Martín-Gómez et al., 2024; Zhou et al., 2020). As outlined in the Introduction, the present study defines the GBSC as an environmental specialization of the BSC, and both its field validation and the modeling of the cause-and-effect relationships between its perspectives are left to future research.

2.3. The four perspectives of the green balanced scorecard

In this study, the Green Balanced Scorecard (GBSC) is considered an approach for integrating environmental objectives and indicators into the four principal BSC perspectives, such that environmental considerations are incorporated into the organization's financial, customer, internal-process, and learning and growth logic and performance assessment extends beyond an exclusive focus on economic outcomes. The operational boundary applied to distinguish environmental from social indicators, and the unit of measurement of the indicators lying closest to that boundary, are specified in Table 3, Panel B, in Section 3.2.1, which distinguishes environmental outcome indicators from environmental enabling indicators.

2.3.1. Green financial perspective

This perspective explains the relationship between environmental activities and their financial consequences within the framework of Environmental Management Accounting (EMA), which expresses the environmental pillar of the TBL in monetary terms within the BSC financial logic. By identifying and allocating environmental costs and benefits to products and processes, EMA provides the information required for managerial decision-making and improved environmental performance. In addition, combining EMA with environmental knowledge management and top-management support can strengthen organizational environmental performance (Bresciani et al., 2023; Stanescu et al., 2021). The indicators in this perspective relate to environmental costs and benefits, environmental risks and liabilities, and green investments and improvement actions. The accounting operationalization of each financial indicator, including its measurement

Uncorrected proof

formula and the disclosure framework under which it is reported, is specified in Table 3, Panel A, in Section 3.2.1.

2.3.2. Green customer perspective

This perspective focuses on customers' perceptions of and responses to the organization's environmental products and actions, reflecting stakeholder expectations and the organization's environmental legitimacy. Research indicates that green product innovation can enhance brand value when accompanied by ethical conduct and organizational transparency, whereas sustainability claims alone do not necessarily lead to sales success (Morea et al., 2023; Van Doorn et al., 2021). Accordingly, the indicators in this perspective relate to environmental perceptions of the product, trust and the brand's ethical credibility, and market responses to products with sustainability attributes.

2.3.3. Green internal-process perspective

This perspective focuses on operational processes and the value chain that affect resource consumption and pollutant emissions, which the NRBV identifies as the locus of pollution prevention and product stewardship capabilities. Studies have shown that energy management, carbon-emission reduction, the use of recycled materials, green supplier selection and integration, and the application of circular-economy principles are among the most important factors for improving environmental performance (Abuzaid et al., 2024; Bendig et al., 2023; Li and Zhong, 2024; Martín-Gómez et al., 2024; Panagiotopoulou et al., 2022). On this basis, the indicators in this perspective include energy and emissions management, material consumption and recycling, green supplier management, and the application of circular-economy principles in design and operations.

2.3.4. Green learning and growth perspective

This perspective forms the foundation for developing an organization's green capabilities and corresponds directly to the NRBV logic of absorptive capacity and sustainable competitive advantage. Managers' environmental competencies, absorptive capacity for environmental knowledge, green training and culture, green knowledge management and green technological innovation are among the most important factors in strengthening organizational environmental performance (Dzhengiz and Niesten, 2020; Khan et al., 2024; Sun et al., 2024). Accordingly, the indicators in this perspective relate to environmental capabilities and absorptive capacity, green training and culture, green knowledge management, and green technological innovation.

2.4. Empirical background

Uncorrected proof

Empirical studies indicate that environmental indicators across all four GBSC perspectives are significantly associated with organizational performance outcomes. In the green financial perspective, Xing et al. (2025), analyzing more than 6,500 observations from polluting-industry firms listed on China's stock exchanges, found that environmental violations increase the likelihood of financial distress and that weak corporate governance intensifies this effect. Bresciani et al. (2023) also confirmed a positive relationship among Environmental Management Accounting (EMA), environmental knowledge management, and financial environmental performance. This evidence indicates that environmental costs, penalties, and green investments are among the most important financial indicators related to green performance.

In the green customer perspective, Van Doorn et al. (2021) found that sustainability claims in new products do not, by themselves, guarantee sales success and that corporate social responsibility (CSR) credibility plays a decisive role in market response. Morea et al. (2023) also found that environmental product innovation affects brand value only through customer trust and ethical perceptions. Accordingly, organizational trust and ethical credibility, alongside the product's green attributes, are key elements of this perspective.

In the green internal-process perspective, Panagiotopoulou et al. (2022) showed that energy management and carbon-emission reduction play a fundamental role in improving the environmental performance of the manufacturing sector. In the supply-chain context, Abuzaid et al. (2024) confirmed the positive effect of green supplier selection, while Li and Zhong (2024) confirmed the effective role of green supplier integration in environmental performance. Martín-Gómez et al. (2024) also demonstrated the application of circular-economy principles within the BSC framework in the plastic-packaging industry. These studies emphasize the importance of energy management, pollutant emissions, recycling, and green supply-chain management.

In the green learning and growth perspective, Sun et al. (2024) showed that green training improves employees' green performance through green organizational culture and green work ethics. Dzhengiz and Niesten (2020) also emphasized the role of absorptive capacity for environmental knowledge in improving performance. In addition, Khan et al. (2024) showed that green knowledge management and green technological innovation have positive and significant effects on organizational sustainable performance, while Amoako et al. (2023) emphasized the role of internal auditing and sustainability auditing in strengthening control and assurance mechanisms.

Uncorrected proof

2.5. Methodological background: application of Fuzzy Delphi

Environmental performance measurement involves linguistic uncertainty in expert judgments, and classical Delphi methods have limited ability to manage this uncertainty. By using Triangular Fuzzy Numbers, the Fuzzy Delphi Method enables the calculation of threshold values, consensus levels, and defuzzification scores, thereby increasing the precision of indicator screening (Ishikawa et al., 1993; Wu and Fang, 2011). Systematic reviews indicate that many studies on sustainable BSC and green manufacturing have relied on secondary data and have made less use of structured expert judgment (Bendig et al., 2023; Jassem et al., 2022). By contrast, Delphi-based studies have shown that expert participation can improve the content validity of environmental indicators (Bouzon et al., 2016).

2.6. Summary and research gap

The literature review indicates that although considerable empirical evidence exists on the positive effects of environmental indicators within each of the four BSC perspectives, these indicators have not yet been integrated into a coherent four-perspective framework based on expert consensus. The existing literature also faces three main gaps: the absence of theoretical integration of BSC, TBL, and NRBV together with stakeholder and legitimacy perspectives in the form of operational GBSC indicators; the limited use of fuzzy consensus-based approaches to manage expert uncertainty; and the lack of an integrated set of environmental indicators that simultaneously covers all four BSC perspectives. By employing document analysis and a two-stage Fuzzy Delphi process, the present study seeks to address these theoretical, methodological, and applied gaps.

2.7. Research framework

2.7.1. Theoretical foundations of the GBSC conceptual framework

As set out in Section 2.1, the study's conceptual framework rests on five theoretical foundations: the Balanced Scorecard for the four-perspective structure and the translation of strategy into measures (Kaplan and Norton, 1992); the Triple Bottom Line for the focus on the environmental dimension and the distinction between "green" and "sustainable" (Elkington, 1997); the Natural Resource-Based View for the role of green capabilities and knowledge in the development of environmental performance (Hart, 1995); and stakeholder theory (Freeman, 1984) and legitimacy theory (Suchman, 1995) for environmental accountability and the maintenance of social acceptance of green actions.

uncorrected proof

These theoretical foundations provide the framework for extracting and validating GBSC indicators, while the Fuzzy Delphi Method is used as the operational mechanism for assessing expert consensus and screening the indicators.

The correspondence between the five theoretical foundations, the four GBSC perspectives and the resulting indicators, together with the criterion applied in each case, is summarized in Table 1.

Table 1. Mapping of the theoretical foundations to the GBSC perspectives and representative indicators

Theoretical foundation	**Core construct**	**Mapping criterion: an indicator is assigned to this foundation when it**	**Perspective**	**Representative indicators**	**Supporting sources**
Balanced Scorecard	Translation of strategy into a balanced set of measures	contributes to the four-perspective architecture as a whole	All four	Structure of the 29 final indicators	Kaplan and Norton (1992); Jassem et al. (2022)
Triple Bottom Line	Separation of the environmen-tal pillar from the economic and social ones	measures an environmental outcome in physical or monetary units (emissions, energy, waste, environmental cost or revenue)	F and IP	F1-F7, F9, IP2, IP7, IP8	Elkington (1997); Jassem et al. (2022)
NRBV – pollution Prevention	Reduction of emissions at source rather than end-of-pipe	reduces emissions, effluents or hazardous inputs at the point of production	IP	IP1, IP2, IP6, IP7, IP8	Hart (1995); Panagioto-poulou et al. (2022)
NRBV – product	Management of the product	extends beyond the firm boundary to	IP and C	IP4, IP10, IP12, IP15,	Hart (1995);

uncorrected proof

Stewardship	life cycle and the supply chain	suppliers, materials, packaging, reuse or recycling		IP16, C2	Abuzaid et al. (2024); Li and Zhong (2024)
NRBV – absorptive capacity and capability building	Accumulation of green knowledge and capability as a source of competitive advantage	builds organizational capability (knowledge, skills, R&D, innovation) rather than measuring an environmental outcome	LG	LG1, LG2, LG4, LG5, LG7, LG8, LG13	Hart (1995); Dzhengiz and Niesten (2020); Khan et al. (2024)
Stakeholder theory	Responsiveness to the expecta-tions of identifiable external parties	captures the response or expectation of an identifiable external party (customer or supplier)	C and IP	C1, C7, IP12	Freeman (1984); Van Doorn et al. (2021); Morea et al. (2023)
Legitimacy theory	Environmental legitimacy, corporate image and regulatory compliance	concerns corporate image, reputation, disclosure or regulatory compliance rather than a physical environmental outcome	C and F	C3, F3	Suchman (1995); Ha (2022); Pajuelo-Moreno et al. (2024)

Note: The mapping criteria in column three were applied to the indicator definitions given in Appendix B. The assignments indicate the primary theoretical justification for each group of indicators and are not mutually exclusive: the five foundations are complementary lenses, and several indicators are informed by more than one of them. For example, the learning and growth indicators also fall within the environmental pillar of the Triple Bottom Line, but their primary justification is the NRBV, because they measure the accumulation of green capability rather than an environmental outcome. Conversely, the emission and waste indicators are grounded in both the Triple Bottom Line, which identifies them as environmental-pillar outcomes, and the NRBV, which identifies pollution prevention as the capability that generates them. The table reproduces the attributions already developed in Sections 2.3.1 to 2.3.4 and 2.7.1 and introduces no new theoretical claim; indicators C4, C5 and C6, which operationalized the legitimacy

Uncorrected proof

dimension most directly, did not reach the consensus threshold and were removed during the Delphi process, as discussed in Section 5.2.2.

2.7.2. Two-Stage Research Process

The study was conducted in two stages. In the first stage, document analysis of 97 scientific sources yielded 76 initial components, which were reduced to 47 initial indicators and classified into the green financial, customer, internal-process, and learning and growth perspectives. In the second stage, the indicators were evaluated using a two-round Fuzzy Delphi Method with 15 experts in the first round and 10 experts in the second. Indicators were accepted according to threshold value ≤ 0.20, consensus level $\geq$ 80% (the ceiling of 0.75N), and defuzzification score ≥ 0.50. The research process and final conceptual framework are presented in Figure 1.

uncorrected proof

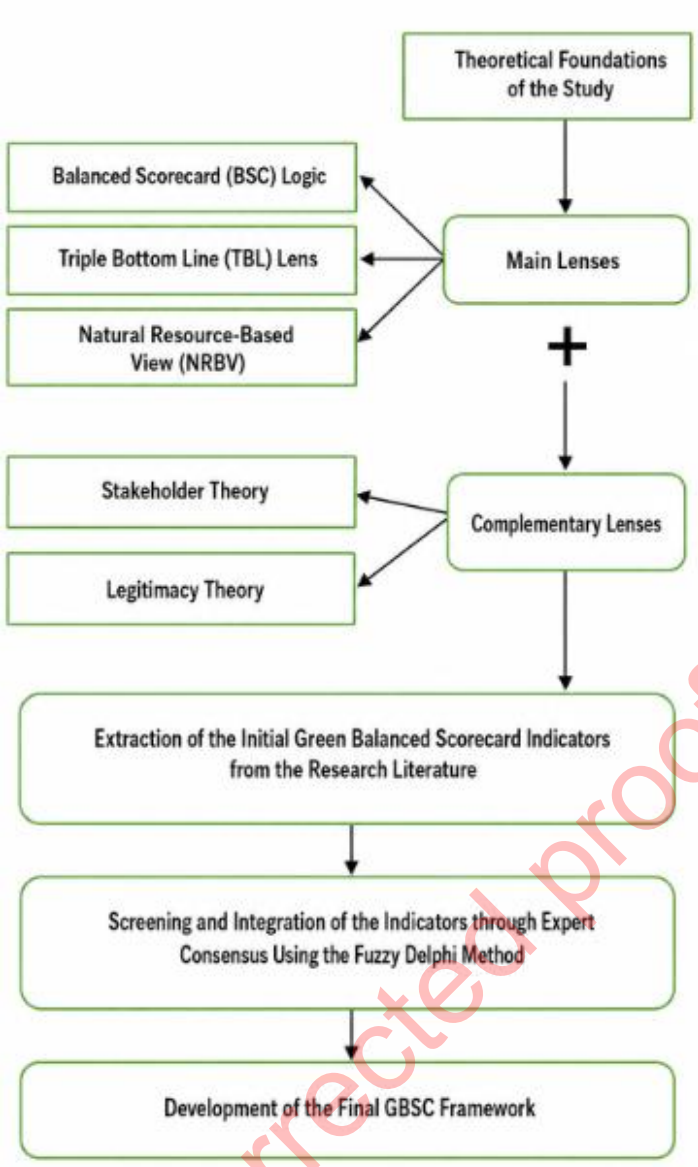


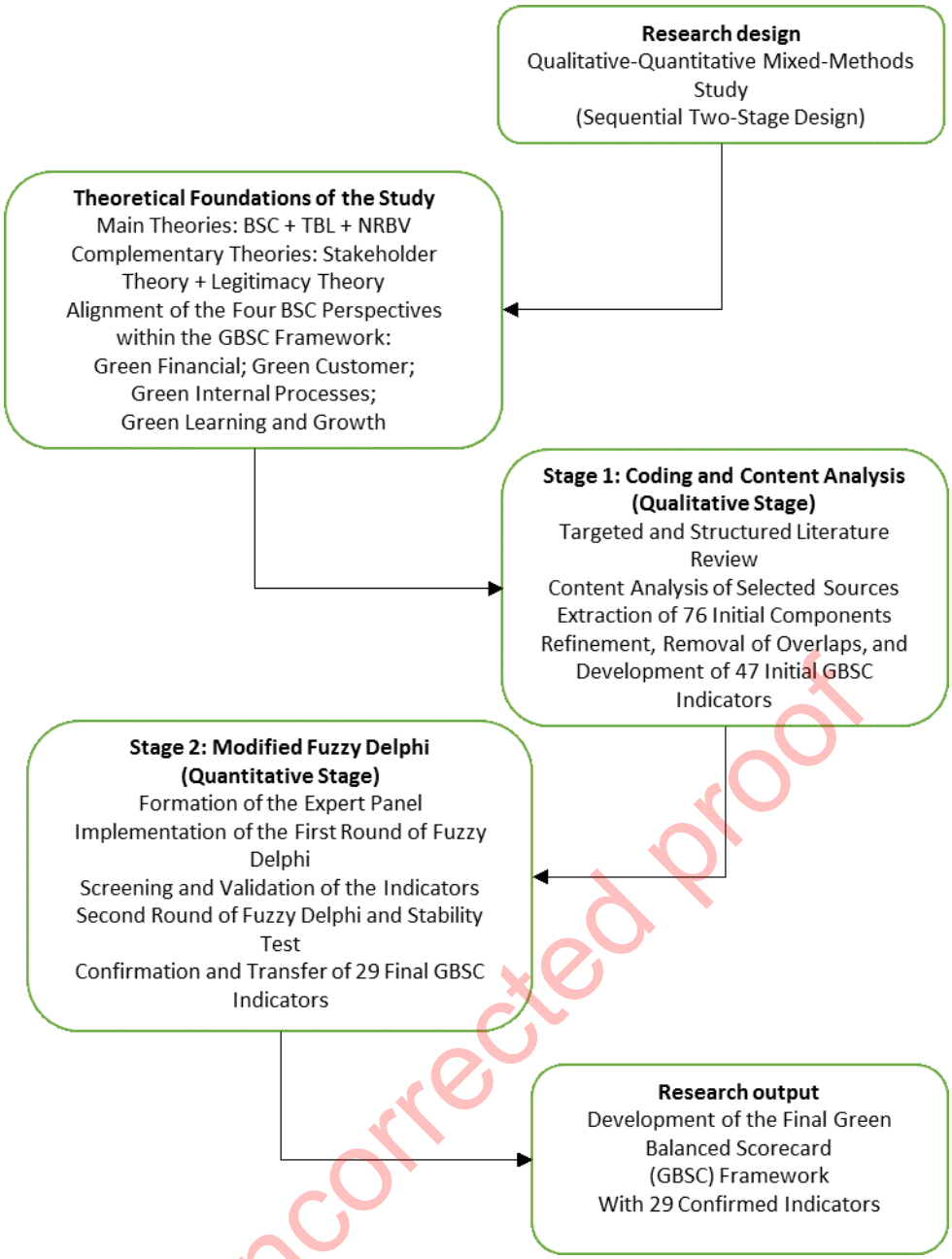
Figure 1. Research framework and the development path of GBSC indicators

uncorrected proof

3. Methodology

3.1. Research design

In terms of purpose, the present study is applied research, and in terms of the nature of the data, it is a qualitative-quantitative mixed-methods study implemented as a sequential two-stage design. An overview of the research implementation stages is presented in Figure 2.



uncorrected proof

Figure 2. Research implementation method

In the first stage, the initial GBSC indicators were extracted through targeted document analysis. In the second stage, a modified Fuzzy Delphi Method was applied in two rounds to screen and validate the indicators, to compare expert responses between rounds, and to achieve expert consensus regarding their importance and appropriateness. The modified Fuzzy Delphi Method was selected because of its capacity to reduce ambiguity arising from experts' subjective judgments, convert qualitative opinions into quantitative values, and achieve consensus on the importance and appropriateness of the indicators.

3.2. Document analysis

In the first stage, to identify the initial GBSC indicators, relevant scientific sources (97 in total) were reviewed and analyzed from reputable international databases, including Scopus, Web of Science, ScienceDirect, Emerald Insight, SpringerLink, and Wiley Online Library. The study's green components were extracted through a targeted and structured literature review. Following structured content analysis, the indicators were classified into the financial, customer, internal-process, and learning and growth perspectives.

The search protocol is summarized in Table 2, and the flow from the included sources to the final indicator set is shown in Figure 3.

Table 2. Search protocol for the document analysis

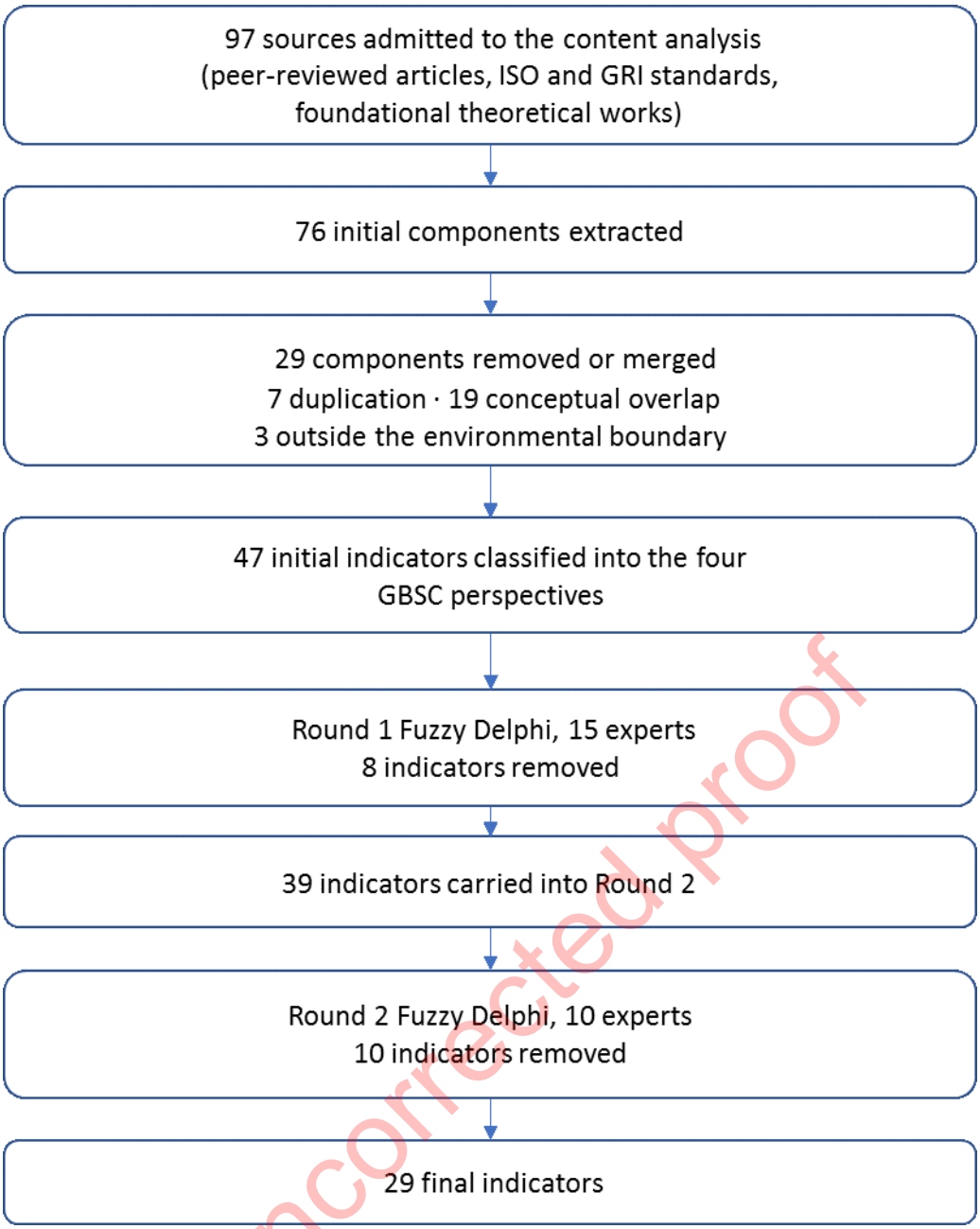
Element	**Specification**
Databases searched	Scopus; Web of Science; ScienceDirect; Emerald Insight; SpringerLink; Wiley Online Library
Search terms	("green balanced scorecard" OR "sustainability balanced scorecard" OR "balanced scorecard") AND ("environmental performance" OR "environmental indicator*" OR "green performance" OR "sustainability indicator*" OR "performance measurement"). Supplementary searches combined each of the four perspectives with "green" or "environmental", for example "green supply chain", "green training", "environmental management accounting", "green brand".
Time horizon	Priority given to studies published from 2020 onwards; earlier works retained only where they provide foundational theory or established indicator definitions.
Language	English
Document types	Peer-reviewed journal articles; international standards and reference documents issued by ISO, GRI, the IFRS Foundation and the United Nations.

Uncorrected proof

Inclusion criteria	(1) direct relevance to the GBSC or to one of its four perspectives; (2) provision of an environmental performance indicator, criterion or measure; (3) publication in a peer-reviewed journal or issuance by a recognized standard-setting body; (4) classical works admitted only where they fulfill a foundational theoretical role
Exclusion criteria	(1) absence of any operational environmental indicator; (2) focus confined to the social or economic pillar of sustainability without an environmental component; (3) non-peer-reviewed outlets; (4) duplication of an already included source
Sources included in the content analysis	97
Date of the final search	November 2024, with the literature updated during revision in November 2026

Note: The review was a structured, purposive literature search designed to identify candidate indicators, not an exhaustive systematic review; accordingly, the number of records initially retrieved from each database was not logged, and the flow reported in Figure 3 begins with the 97 sources admitted to the content analysis. The reduction from the 76 extracted components to the 47 initial indicators is documented in Appendix C.

uncorrected proof



uncorrected proof

Figure 3. Flow of sources, components and indicators from the document analysis to the final indicator set

Figure 3 shows the stages used to extract the initial GBSC indicators from the research literature. In addition to scientific articles, international standards and reference documents such as ISO 50001, ISO 14001, and the GRI standards (303, 305, and 306) were used as supplementary and authoritative sources. This stage yielded 76 initial components; after duplicate and overlapping items were removed, 47 initial indicators were developed across the four GBSC perspectives. The coding was performed by the first author and reviewed by the supervising author; the decision rules applied and the

number of components removed under each are reported in Appendix C. The complete list of the 47 extracted indicators, together with the supporting sources for each indicator, is presented in Appendix B. For brevity, indicators in the financial perspective are denoted by F, those in the customer perspective by C, those in the internal-process perspective by IP, and those in the learning and growth perspective by LG.

The indicators extracted during the document-analysis stage formed the basis for designing the first-round Fuzzy Delphi questionnaire.

3.2.1. Operational definition of the indicators: accounting treatment and environmental boundary

The operational definition of the indicators is specified in Table 3. Panel A maps each financial indicator to a measurement formula, to the accounting treatment through which it is generated and to the disclosure framework under which it is reported. Panel B states the unit of measurement and the environmental boundary of the six indicators lying closest to the margin with corporate social responsibility.

Table 3. Operational definition of the indicators: accounting treatment and environmental boundary

Panel A. Accounting operationalization of the green financial indicators				
Code	**Indicator**	**Measurement formula**	**Accounting treatment or technique**	**Related standard or reference**
F1	Green product and service revenue	Green revenue / total revenue	Disaggregation of revenue	IFRS 15; IFRS S2
F2	Green ancillary revenues	Revenue from recovered materials / other operating income	Other operating income	GRI 306
F3	Environmental non-compliance costs	Fines and penalties recognized in the period,	Provisions and contingent liabilities	IAS 37; GRI 2-27

uncorrected proof

		together with any related provisions		
F4	Environmental operating costs	Prevention, control and disposal costs / total operating costs	Environmental cost categories in management accounting	ISO 14051; UN DSD EMA guidance
F5	Resource-consumption optimization	Monetary value of energy and material savings	Material flow cost accounting; cost of goods sold	ISO 14051; ISO 50001
F6	Waste-cost management	Material-loss cost / total production cost	Material flow cost accounting; inventory valuation	IAS 2; ISO 14051; GRI 306
F7	Investment in sustainable technologies	Environmental capital expenditure / total capital expenditure	Property, plant and equipment; intangible assets	IAS 16; IAS 38; IFRS S2
F9	Return on green investment	Savings and income from green investment / green investment base	Capital budgeting; internal carbon price	IFRS S2

Panel B. Environmental boundary of the indicators at the margin with corporate social responsibility

Code	Class	Indicator	Unit of measurement	Why environmental and not social

uncorrected proof

IP1	Outcome	Green manufacturing process	Energy, material and waste intensity per unit of output	Occupational safety and quality control were excluded at the extraction stage
IP12	Enabling	Assessment of supplier environmental performance	Share of suppliers assessed against environmental criteria	Assessment criteria are environmental only; labor and human rights criteria are excluded
C2	Outcome	Sustainable product design and packaging	Share of packaging that is recyclable or bio-based	Measured by material composition, not by consumer welfare
LG1	Enabling	Development of employees' knowledge of the organization's environmental impacts and environmental management procedures	Share of staff passing an environmental-awareness assessment	Restricted to knowledge of the firm's environmental impacts and environmental management procedures; excludes general human resource development
LG2	Enabling	Employees' environment-specific technical skills	Share of staff certified in environment-specific technical competences	Environment-specific technical competences only, such as energy auditing, waste segregation and life-cycle assessment
LG13	Enabling	Linking employee incentives to	Share of variable pay or recognition	A management control mechanism linked to environmental outcomes,

uncorrected proof

		environmental objectives	schemes linked to environmental KPIs	not a general employee engagement measure

Note: Panel A expresses each financial indicator in terms of line items and cost categories already recognized in financial and management accounting, so that it can be populated from the entity's existing accounting records. Indicators IP4, IP8 and IP15 supply the physical quantity data required to operate material flow cost accounting for F5 and F6, which is the point at which the physical and monetary information systems of environmental management accounting converge .Panel B applies the admission criterion of the GBSC, which recognizes two classes of indicator. An outcome indicator measures an environmental flow or impact in a physical environmental unit or its monetary equivalent. An enabling indicator measures an organizational capability, process or control mechanism defined exclusively by reference to the environmental domain and directed at improving environmental performance. This class comprises the whole of the learning and growth perspective, which in the Balanced Scorecard is by construction enabling rather than outcome-based, together with those internal-process indicators that measure a control mechanism rather than a physical flow. Both classes are confined to the environmental pillar: indicators relating to occupational health and safety, labor rights, community engagement, philanthropy, and diversity and inclusion are excluded from both and belong to the social pillar of the SBSC. The indicator codes correspond to those listed in Appendix B.

3.3. Expert panel and rationale for its disciplinary composition

The validity of Fuzzy Delphi results depends more on the depth of expert knowledge than on the number of respondents (Okoli and Pawlowski, 2004); accordingly, the panel was selected using explicit criteria. Fifteen experts participated in the first round, and 10 of the first-round experts participated in the second round, consistent with the standard proposed by Okoli and Pawlowski (2004) for homogeneous Delphi panels (10 to 18 experts).

Purposive, non-random sampling was used based on four criteria: (1) an academic rank of at least assistant professor, or an adjunct lecturer holding a doctoral degree; (2) at least 4 years of teaching experience; (3) documented research experience in green accounting, sustainability management, or natural resource economics; and (4) self-assessed familiarity with the GBSC at level 4 or 5 on a five-point scale.

The panel was designed with three complementary disciplinary groups consistent with the interdisciplinary nature of the GBSC: accounting (5 experts) for the financial perspective and part of the internal-process perspective; management (5 experts) for the customer and learning and growth perspectives; and agricultural economics (5 experts) for the dimensions of sustainability, natural resources, the environment, and policy-making.

uncorrected proof

The scientific rationale for including agricultural economics was based on the international JEL classification. Under this classification, agricultural economics, natural resource economics, and environmental economics fall within subject category Q (Agricultural and Natural Resource Economics; Environmental and Ecological Economics), reflecting the thematic connection and research overlap among these fields in the economics literature. Detailed demographic characteristics of the 15 experts are presented in Table 4.

Table 4. Demographic characteristics of the 15 experts on the Fuzzy Delphi panel

Variable	Category	Frequency	Percentage
Gender	Male	12	80%
	Female	3	20%
	Total	15	100%
Age	30-39 years	4	26.7%
	40-49 years	3	20%
	50-59 years	5	33.3%
	60 years and above	3	20%
	Total	15	100%
Academic Rank	Full Professor	2	13.3%
	Associate Professor	5	33.3%
	Assistant Professor	5	33.3%
	Adjunct Lecturer (PhD)	3	20%
	Total	15	100%
Teaching Experience	4-9 years	5	33.3%
	10-19 years	4	26.7%
	20 years and above	6	40%
	Total	15	100%
Field of Study	Agricultural Economics	5	33.3%
	Accounting	5	33.3%
	Management	5	33.3%
	Total	15	100%
Level of Familiarity with GBSC	Good	3	20%
	Very Good	12	80%
	Total	15	100%

uncorrected proof

3.4. Questionnaire design, validity, and reliability

3.4.1. Questionnaire structure and rationale for changing the scale between rounds

The Fuzzy Delphi questionnaires were designed in two rounds (the first-round Delphi questionnaire is provided in Appendix A). The first round comprised 47 indicators across the four GBSC perspectives and used a five-point Likert scale to capture the full range of experts' linguistic judgments. The five-point scale was selected for the first round in accordance with the recommendation of Hsu and Yang (2000) and to align with the standard Fuzzy Delphi literature, because a five-level scale in the exploratory phase enables representation of the full range of experts' linguistic judgments, including genuine indifference or non-applicability for indicators outside an expert's immediate domain (Nadler et al., 2015). In the analysis the neutral category was coded as the midpoint of the normalized scale and entered the fuzzy aggregation as a neutral judgment rather than being excluded as missing data; it was removed in the confirmatory round so that no response would be recorded as neutral by default.

The second-round questionnaire included the 39 indicators remaining after first-round screening and used an even four-point Likert scale (with no neutral option). The deliberate removal of the midpoint in the second round was based on two theoretical considerations: (a) reducing central-tendency bias, which in consensus-based studies, particularly at the final judgment stage, weakens the reliability of ranking (Chomeya, 2010; Chyung et al., 2017); (b) requiring experts to take an explicit position in the confirmatory round so that acceptable and unacceptable indicators could be distinguished more clearly.

To address the heterogeneity resulting from the two different scales, responses from both rounds were transformed to the common interval [0, 1] using Min-Max Normalization before any fuzzy calculation. The relationship used was as follows:

$$x_norm = (x - x_min) / (x_max - x_min)$$

where x is the raw response value on the Likert scale, and x_min and x_max are, respectively, the lower and upper bounds of the scale in each round. The normalized values were then converted, according to the conversion table (Table 6), into Triangular Fuzzy Numbers (TFNs) to ensure comparability between the two rounds in fuzzy space. The sensitivity of the defuzzified scores to this conversion is reported in Table 10. The

uncorrected proof

characteristics of the two versions of the Fuzzy Delphi questionnaire are presented in Table 5.

Table 5. Characteristics of the two versions of the Fuzzy Delphi questionnaire

Round	Number of Indicators	Measurement Scale	Number of Respondents	Response Rate
First (Exploratory)	47	5-point Likert scale	15	100%
Second (Confirmatory)	39	4-point Likert scale	10	66.7%

3.4.2. Instrument validity

Questionnaire validity was established through two complementary routes:

(a) Face validity: Before the survey, a copy of the questionnaire was provided to three experts from the panel to confirm the readability and clarity of the items and their apparent relevance to the research objective.

(b) Content validity: Content validity was established through support from the literature, ISO and GRI standards, and the expert-consensus process in the Fuzzy Delphi procedure.

3.4.3. Between-round comparison of expert responses and assessment of instrument reliability

In Fuzzy Delphi studies conducted with small, specialized panels, Cronbach's alpha is not considered an appropriate measure of reliability. For reasons such as the independence of the indicators (measurement of distinct concepts), the small panel size (10 to 18 experts), and the consensus-oriented nature of the method, Cronbach's alpha is not suitable for assessing reliability. Accordingly, the reliability of the study is evaluated by examining the consistency of responses across successive rounds (Dajani et al., 1979; Von Der Gracht, 2012).

On this basis, the consistency of expert responses between the two Delphi rounds in the present study was examined descriptively using the nonparametric paired Wilcoxon signed-rank test. This test performs a paired comparison of each expert's responses in the first and second rounds, and its null hypothesis is based on the absence of a significant difference between the two rounds. Failure to reject the null hypothesis at the 0.05 significance level is interpreted as an indication that expert responses did not shift

Uncorrected proof

systematically between the two rounds. The primary stability criterion, the stability percentage for each indicator (S_i), is defined as follows:

$$S_i = \frac{n(|x_{i1} - x_{i2}| \leq 0.25)}{N} \times 100$$

where x_{i1} and x_{i2} denote expert i's normalized responses in the first and second rounds respectively, and N is the number of second-round experts.

3.5. Fuzzy analysis and indicator acceptance criteria

3.5.1. Conversion of linguistic variables to triangular fuzzy numbers

After the Likert values had been standardized to the common interval [0, 1] in the preceding section, each linguistic expert response was converted into a Triangular Fuzzy Number (TFN). A triangular fuzzy number is a three-component representation that expresses the inherent uncertainty of human judgment within a meaningful interval and is defined as follows:

$$\tilde{A} = (l, m, u), \text{ subject to } l \leq m \leq u$$

where l denotes the lower bound, m the most plausible value, and u the upper bound of the expert's response. The corresponding membership function is defined as

$$\begin{aligned} mu(x) &= 0 && \text{for } x < l \\ mu(x) &= (x - l) / (m - l) && \text{for } l <= x <= m \\ mu(x) &= (u - x) / (u - m) && \text{for } m <= x <= u \\ mu(x) &= 0 && \text{for } x > u \end{aligned}$$

so that l and u delimit the support of the judgment and m is the value with full membership. Table 6 presents the conversion of the linguistic variables of both Delphi rounds into TFN values accepted in the standard Fuzzy Delphi literature (Ishikawa et al., 1993; Hsu and Yang, 2000; Wu and Fang, 2011) .

uncorrected proof

Table 6. Conversion of linguistic variables into triangular fuzzy numbers (TFNs) in both Delphi rounds

Round	Scale	Linguistic variable	Code	Likert	Normalized	l	m	u	Source of mapping
1	5-point	Strongly disagree	SD	1	0.00	0.00	0.10	0.25	Ishikawa et al. (1993); Hsu and Yang (2000); Wu
1	5-point	Disagree	D	2	0.25	0.10	0.25	0.50	
1	5-point	No opinion	N	3	0.50	0.25	0.50	0.75	

1	5-point	Agree	A	4	0.75	0.50	0.75	0.90	and Fang (2011)
1	5-point	Strongly agree	SA	5	1.00	0.75	0.90	1.00	
2	4-point	Strongly disagree	SD	1	0.00	0.00	0.00	0.33	Ishikawa et al. (1993), adapted to an even scale
2	4-point	Disagree	D	2	0.33	0.00	0.33	0.67	
2	4-point	Agree	A	3	0.67	0.33	0.67	1.00	
2	4-point	Strongly agree	SA	4	1.00	0.67	1.00	1.00	

Note: Raw Likert values were first normalised to the common interval [0, 1] using Min-Max normalisation and then converted into triangular fuzzy numbers using the mapping corresponding to that round. The robustness of the results to this round-specific conversion is examined in Table 10.

Worked example (indicator LG4, Round 2, N = 10): all ten experts selected "Strongly agree", so each response maps to (0.67, 1.00, 1.00). The group fuzzy mean is therefore (0.670, 1.000, 1.000); the centroid defuzzified score is (0.670 + 1.000 + 1.000) / 3 = 0.890; the geometric distance of every expert from the group mean is 0.000; and the consensus percentage is 100%. The indicator satisfies all three acceptance criteria and is retained, as reported in Table 9.

uncorrected proof

3.5.2. Defuzzification using the centroid method

For each indicator, the experts' group fuzzy mean was first calculated by component-wise averaging of the three vertices of the triangular fuzzy number:

$$\bar{l} = (1 / n) \times (l_1 + l_2 + \ldots + l_n)$$

$$\bar{m} = (1 / n) \times (m_1 + m_2 + \ldots + m_n)$$

$$\bar{u} = (1 / n) \times (u_1 + u_2 + \ldots + u_n)$$

where n is the number of responding experts and (l_i, m_i, u_i) is the TFN corresponding to expert i's response. Accordingly, the group fuzzy mean is obtained as $\bar{\bar{A}} = (\bar{l}, \bar{m}, \bar{u})$.

In the next step, to make the fuzzy values comparable and to rank the indicators quantitatively, each group fuzzy number was converted into an equivalent crisp value using the centroid method, also known in the fuzzy literature as the Simple Centroid method (Yager, 1981; Chen and Hwang, 1992):

$$DF_i = (\bar{l} + \bar{m} + \bar{u}) / 3$$

The DF_i value, which lies within the interval [0, 1], reflects the group's mean importance level for indicator i and provides the basis for comparison with the acceptance threshold.

3.5.3. Consensus measurement based on geometric distance between triangular fuzzy numbers

The degree of agreement between each expert and the group mean was calculated using geometric distance based on the vertex method. This distance is a standard and widely used measure in the fuzzy decision-making literature and is defined as follows:

$$d_i = \sqrt{\{ (1/3) \times [(l_i - \bar{l})^2 + (m_i - \bar{m})^2 + (u_i - \bar{u})^2] \}}$$

where d_i is the geometric distance between expert i's judgment and the group mean. The above relationship defines a true metric on the space of triangular fuzzy numbers; that is, it satisfies the three fundamental properties of non-negativity, symmetry, and the triangle inequality and is therefore a valid criterion for measuring the deviation of each expert's judgment from the collective judgment.

Based on the standard literature (Bouzon et al., 2016; Cheng and Lin, 2002; Cheng and Mon, 1994; Ishikawa et al., 1993), expert i is considered to be "in agreement with the group" if the geometric distance between that expert's judgment and the group mean does not exceed the threshold of 0.20; that is, $d_i \leq 0.20$. The consensus percentage for each indicator is then obtained from the following relationship:

C_i = (number of experts for whom $d_i \leq 0.20$ / total number of experts) × 100

In addition, the mean geometric distance for each indicator ($\bar{d}_i$) is obtained by averaging the d_i values of all responding experts for that indicator.

3.5.4. Three indicator acceptance criteria

In accordance with the standard Fuzzy Delphi Method (FDM) framework and to ensure transparency in the consensus rules, three acceptance criteria were specified a priori before the analysis was conducted. Indicator i is retained at the end of each Delphi round only if all three of the following conditions are simultaneously satisfied:

(1) Mean geometric distance: $\bar{d}_i \leq 0.20$

(2) Group consensus percentage: $C_i \geq \lceil 0.75N \rceil$ (12 of 15 experts in Round 1; 8 of 10 in Round 2, i.e. 80%)

(3) Defuzzification score: $DF_i \geq 0.50$

A summary of these criteria, together with the theoretical basis for each, is presented in Table 7.

Table 7. Three indicator acceptance criteria in the Fuzzy Delphi method

Row	Acceptance Criterion	Mathematical Symbol	Reference Threshold	Methodological Meaning
1	Mean Geometric Distance (Threshold)	$\bar{d}_i \leq 0.20$	0.20	Homogeneity in the distribution of expert opinions and low dispersion of judgments

uncorrected proof

2	Group Consensus Percentage (Group Consensus)	$C_i \geq [0.75N]$	75% nominal; 80% effective	Proportion of experts whose geometric distance from the group mean is below the 0.20 threshold
3	Defuzzification Score (Defuzzified Score)	$DF_i \geq 0.50$	0.50	Importance level of the group fuzzy mean; the midpoint of the interval [0, 1] as the minimum threshold for indicator acceptance

Sources: Threshold: Ishikawa et al. (1993); Cheng and Mon (1994);

Cheng and Lin (2002); Bouzon et al. (2016)

Consensus: Ishikawa et al. (1993); Murray et al. (1985)

Defuzzification: Ishikawa et al. (1993); Murray et al. (1985)

Note: Indicator i is retained at the end of each Delphi round only if all three conditions are satisfied simultaneously. Because the number of experts is discrete, criterion (2) requires 12 of 15 experts in Round 1 and 8 of 10 experts in Round 2; the effective consensus rate applied in both rounds was therefore 80%, which is more conservative than the nominal 75% benchmark.

3.5.5. Composite ranking index

After completion of the second Delphi round and identification of the set of accepted indicators, a two-dimensional composite index entitled the Composite Ranking Index (CRI) was designed for final ranking and prioritization. This index simultaneously aggregates the two dimensions of importance and consensus into a single crisp value and prevents the ranking from becoming one-dimensional.

Because the two components are measured on different scales, both were transformed to the uniform interval [0, 1] using the Min-Max method before aggregation:

$$\hat{X}_i = (X_i - X\,\min) / (X\,\max - X\,\min) , X \in \{ DF , C \}$$

The CRI for each indicator i was then defined as follows:

$$CRI_i = 0.50 \times \hat{C}_i + 0.50 \times \hat{D}F_i$$

where $\hat{C}_i$ is the normalized consensus percentage in the second Delphi round; $\hat{D}F_i$ is the normalized defuzzification score.

Rationale for the weighting structure: The weighting structure proposed in this study (0.50 for consensus and 0.50 for importance) was designed on the basis of the "consensus-importance" framework in the Fuzzy Delphi literature. Because both components of the CRI are discrete under the present design, ties may occur. Ties were resolved by an explicit hierarchy applied uniformly across all tied bands: first the lower mean geometric distance d-bar reported in Table 9, then the higher between-round agreement index S_i reported in Table 13, and finally alphabetical order of the indicator code.

uncorrected proof

The Friedman test was performed on the Round-2 responses of the n = 10 second-round experts. It was computed separately within each GBSC perspective, with k denoting the number of accepted indicators in that perspective and df = k - 1, and Kendall's coefficient of concordance was reported as the effect size. Mean ranks across all 29 accepted indicators are reported descriptively.

All fuzzy calculations (TFN averaging, centroid defuzzification, and vertex-based geometric distance) were implemented through custom functions in R software (version 4.5.1). The nonparametric tests (paired Wilcoxon and Friedman) were conducted using the stats and DescTools packages, respectively, and data were managed using the dplyr and tidyr packages. Multiplicity-adjusted p-values were obtained with the Benjamini-Hochberg procedure.

3.6. Implementation of the two delphi rounds, screening, and final ranking

The first-round questionnaire was sent to 15 experts, yielding a response rate of 100%. After simultaneous application of the three acceptance criteria-fuzzy threshold, consensus percentage, and defuzzification score-8 of the 47 initial indicators that performed poorly on at least one of the three criteria were removed, and 39 indicators proceeded to the second round. A summary of the statistical feedback for each indicator (including the group fuzzy mean, mean geometric distance, consensus percentage, and the Round-1 retention decision), together with the second-round questionnaire, was then resent to the 15 first-round experts. In the second round, 10 experts returned their responses (a response rate of 66.7%), thereby completing the Delphi feedback cycle in accordance with the standard protocol of Okoli and Pawlowski (2004) and Hasson et al. (2000). Attrition was examined empirically rather than assumed to be benign: Round-1 responses of the 10 Round-2 completers were compared with those of the 5 non-respondents, and the Round-1 defuzzified scores were recomputed on the completers alone. The results are reported in Appendix D.

Second-round screening: After simultaneous application of the three acceptance criteria-fuzzy threshold, consensus percentage, and defuzzification score-10 of the 39 remaining indicators that performed poorly on at least one of the three criteria were removed, and 29 indicators formed the final set.

To describe changes in expert responses between the two rounds, the nonparametric paired Wilcoxon signed-rank test was used. For each indicator, this test compares the paired responses of the same expert across the two rounds. The null hypothesis was specified as follows:

Uncorrected proof

H_0: The median difference in responses between the first and second rounds is equal to zero (no significant difference).

H_1: The median difference in responses between the two rounds differs significantly from zero.

Final ranking of the indicators: The 29 indicators accepted at the end of the second round were ranked using the CRI as defined in Section 3.5.5, and the nonparametric Friedman test was applied within each perspective to the ranks assigned by the second-round respondents. Table 8 summarizes the research implementation stages and the decision criteria at each step.

Table 8. Summary of the research implementation steps in the two stages of document analysis and modified Fuzzy Delphi

Step	Stage	Input	Tool/Test	Acceptance Criterion	Output
1	Targeted Document Analysis	97 scientific sources	Structured content analysis	Topical relevance + scientific support	47 indicators
2	Fuzzy Delphi - Round 1	47 indicators and 15 experts	5-point Likert scale + TFN	$\bar{d}_i \leq 0.20$, C_i ≥ ⌈0.75N⌉, $DF_i \geq 0.50$	39 indicators
3	Fuzzy Delphi - Round 2	39 indicators and 10 experts	4-point Likert scale + TFN	$\bar{d}_i \leq 0.20$, C_i ≥ ⌈0.75N⌉, $DF_i \geq 0.50$	29 accepted indicators
4	Between-round comparison	29 accepted indicators	Paired Wilcoxon; median shift; IQR change	Descriptive only	Between-round agreement 80–100%
5	Final Ranking	29 accepted indicators	CRI; Friedman mean ranks	CRI = 0.50 Ĉ + 0.50 D̂F	Final ranked list of 29 indicators

uncorrected proof

4. Research Findings

4.1. Overview

This section reports the findings from the two rounds of the modified Fuzzy Delphi process. The results are presented in accordance with the methodological stages of the study: first, the first-round results and initial screening of the indicators; then the second-round results and screening of the final indicators; followed by ranking of the accepted indicators based on the CRI. The final four-perspective GBSC framework is then presented. Interpretation and discussion of the findings are provided in Section 5. The

results of the second-round fuzzy analysis are presented in Table 9. In this table, indicators in the financial perspective are denoted by F, customer indicators by C, internal-process indicators by IP, and learning and growth indicators by LG.

Table 9. Results of the second-round Fuzzy Delphi analysis for 39 indicators remaining after round 1

Item	Threshold	Consensus	Fuzzy_L	Fuzzy_M	Fuzzy_U	Defuzzified	Decision
IP10	0.132	90	0.670	1.000	1.000	0.890	Accepted
IP16	0.120	90	0.670	1.000	1.000	0.890	Accepted
LG1	0.084	100	0.670	1.000	1.000	0.890	Accepted
LG4	0.046	100	0.670	1.000	1.000	0.890	Accepted
LG13	0.133	90	0.670	1.000	1.000	0.890	Accepted
IP7	0.090	100	0.636	0.967	1.000	0.867	Accepted
IP15	0.116	100	0.636	0.967	1.000	0.867	Accepted
LG5	0.146	80	0.636	0.967	1.000	0.867	Accepted
IP6	0.140	80	0.602	0.934	1.000	0.845	Accepted
IP9	0.160	70	0.602	0.934	1.000	0.845	Rejected
IP14	0.186	70	0.602	0.934	1.000	0.845	Rejected
LG7	0.111	90	0.602	0.934	1.000	0.845	Accepted
F1	0.097	100	0.602	0.934	1.000	0.845	Accepted
C3	0.110	100	0.602	0.934	1.000	0.845	Accepted
IP3	0.199	70	0.568	0.901	1.000	0.823	Rejected
IP4	0.130	80	0.568	0.901	1.000	0.823	Accepted
IP5	0.178	70	0.568	0.901	1.000	0.823	Rejected
IP8	0.143	80	0.568	0.901	1.000	0.823	Accepted
IP12	0.169	80	0.568	0.901	1.000	0.823	Accepted
LG3	0.182	70	0.568	0.901	1.000	0.823	Rejected
LG6	0.227	60	0.568	0.901	1.000	0.823	Rejected
LG8	0.156	90	0.568	0.901	1.000	0.823	Accepted
LG10	0.166	70	0.568	0.901	1.000	0.823	Rejected
LG12	0.212	70	0.568	0.901	1.000	0.823	Rejected
F4	0.105	100	0.568	0.901	1.000	0.823	Accepted
F5	0.084	90	0.568	0.901	1.000	0.823	Accepted
F7	0.117	80	0.568	0.901	1.000	0.823	Accepted
F9	0.143	80	0.568	0.901	1.000	0.823	Accepted

Uncorrected proof

C1	0.124	80	0.568	0.901	1.000	0.823	Accepted
C2	0.143	80	0.568	0.901	1.000	0.823	Accepted
IP1	0.116	90	0.534	0.868	1.000	0.800	Accepted
LG2	0.165	80	0.534	0.868	1.000	0.800	Accepted
LG9	0.205	60	0.534	0.868	1.000	0.800	Rejected
F6	0.116	100	0.534	0.868	1.000	0.800	Accepted
C7	0.131	80	0.534	0.868	1.000	0.800	Accepted
IP2	0.124	80	0.500	0.835	1.000	0.778	Accepted
F2	0.150	80	0.500	0.835	1.000	0.778	Accepted
F3	0.143	80	0.500	0.835	1.000	0.778	Accepted
C4	0.247	50	0.500	0.835	1.000	0.778	Rejected

Note: Identical Fuzzy_L/M/U triplets across different indicators are an expected consequence of the discreteness of the four-point confirmatory scale combined with the small confirmatory panel (n = 10). With four linguistic levels and ten experts, the component-wise group fuzzy mean can take only a limited number of distinct values, so any two indicators with the same response-frequency vector necessarily yield the same triangular fuzzy number and the same defuzzified score. The indicators are nevertheless distinguished by the mean geometric distance and the consensus percentage reported in the first two columns, which are computed from the full individual-level response distribution. These values are not a data-entry error.

The results in Table 9 show that, of the 39 indicators evaluated in the second round, 29 satisfied all three acceptance criteria (threshold ≤ 0.20, consensus ≥ 80% -the ceiling of 0.75N- and defuzzification ≥ 0.50) and formed the final set. Indicators IP10, IP16, LG1, LG4, and LG13 had the highest defuzzification scores. By contrast, 10 indicators were removed, mainly owing to insufficient consensus or a threshold above the allowable limit. Table 10 reports a sensitivity analysis in which a single symmetric triangular fuzzy conversion with identical support was applied to both rounds; the ordering of the defuzzified scores was unchanged (Spearman's rho = 1.000).

Table 10. Sensitivity of the second-round defuzzified scores to the triangular fuzzy conversion

Specification	TFN conversion applied	Range of defuzzified scores	Spearman's rho of the ordering of the 39 indicators
S0 (baseline)	Table 6, Round 2 mapping	0.778 - 0.890	1.000 (reference)
S1 (symmetric)	(x - 0.25, x, x + 0.25), truncated to [0, 1]	0.792 - 0.917	1.000

uncorrected proof

Note: Under S1 a single triangular fuzzy conversion with identical support was applied to both rounds. Group fuzzy means and centroid defuzzified scores were recomputed from the raw expert responses. The alternative conversion produces a uniform upward shift in the defuzzified scores without altering their relative ordering; the consensus percentages and the acceptance decisions reported in Table 9 are those of the baseline specification.

4.2. First-Round Fuzzy Delphi results

The first-round questionnaire was distributed to 15 experts and collected with a response rate of 100%. For the 47 indicators, the group fuzzy mean, geometric distance, consensus percentage, and defuzzification score (DF_i) were calculated. After the three acceptance criteria were applied, 8 indicators were removed and 39 proceeded to the second round. The accepted indicators comprised 8 financial indicators, 5 customer indicators, 14 internal-process indicators, and 12 learning and growth indicators. The list of removed indicators is presented in Table 11.

Table 11. List of indicators removed in the first round of Fuzzy Delphi

Item	Status
IP11	Removed After Round 1
IP13	Removed After Round 1
LG11	Removed After Round 1
LG14	Removed After Round 1
F8	Removed After Round 1
F10	Removed After Round 1
C5	Removed After Round 1
C6	Removed After Round 1

uncorrected proof

According to Table 11, two indicators from each perspective were removed in the first Delphi round. In the financial perspective, these were "investment in the environmental sector" and "increased profitability through a product-differentiation strategy with a green orientation"; in the customer perspective, "environmental risk management and legal compliance" and "strength of regulatory relationships/compliance with future environmental regulations"; in the internal-process perspective, "use of environmental auditing and assessment to manage and control supplier performance" and "environmentally friendly packaging"; and in the learning and growth perspective, "internal performance evaluation system incorporating environmental factors" and "suggestions for environmental improvement".

4.3. Second-Round Fuzzy Delphi results and second-round screening

A summary of the first-round statistical feedback, together with the second-round questionnaire, was sent to the 15 experts; 10 responded (a response rate of 66.7%). All three disciplines remained represented in the second round, although their proportions changed; the non-response analysis reported in Appendix D shows that this did not produce systematic differences in Round-1 judgments. The results showed that 29 indicators met the three acceptance criteria and were retained in the final set, whereas 10 indicators failed at least one criterion and were removed.

Table 12. List of indicators removed in the second round of Fuzzy Delphi

Item	Threshold	Consensus (%)	Defuzzified
C4	0.247	50	0.778
IP3	0.199	70	0.823
IP5	0.178	70	0.823
IP9	0.160	70	0.845
IP14	0.186	70	0.845
LG3	0.182	70	0.823
LG6	0.227	60	0.823
LG9	0.205	60	0.800
LG10	0.166	70	0.823
LG12	0.212	70	0.823

The accepted indicators comprised 8 financial, 4 customer, 10 internal-process, and 7 learning and growth indicators. Among the removed indicators, C4 showed the weakest performance. The level of agreement between the two Delphi rounds for the accepted indicators is shown in Figure 4.

Uncorrected proof

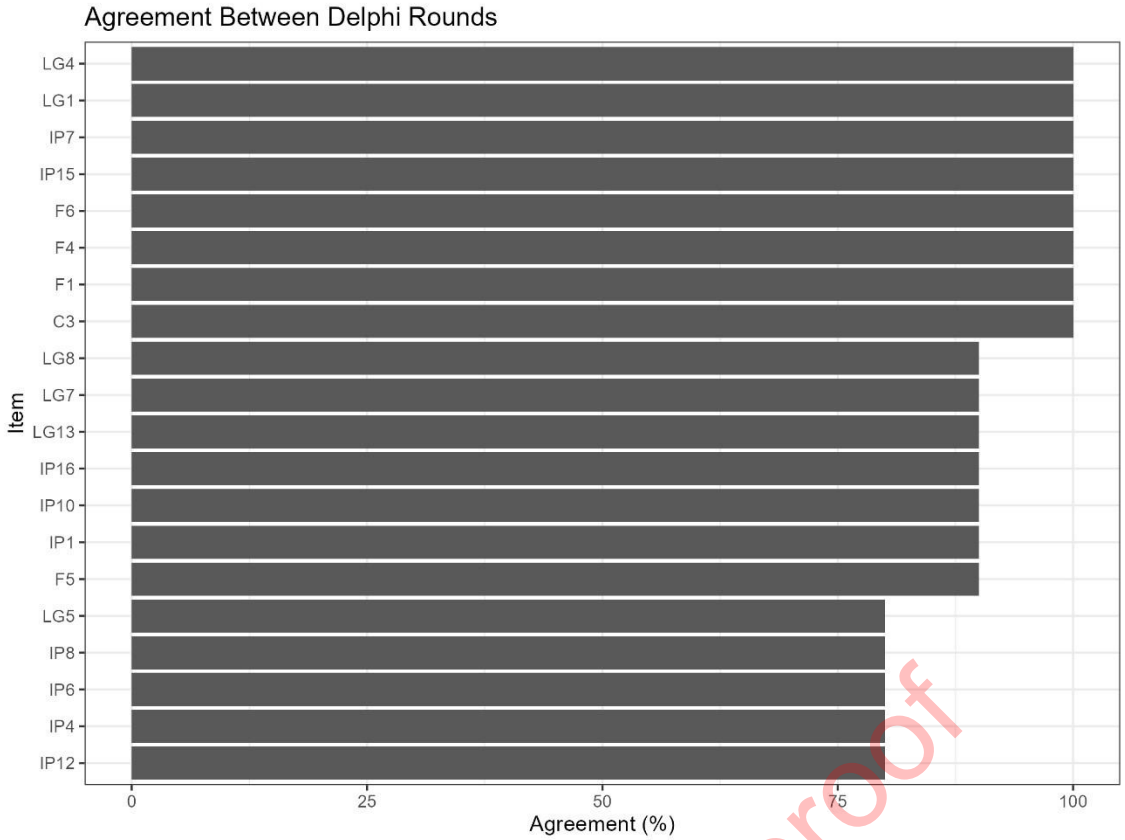


Figure 4. Level of agreement between the two Delphi rounds for the top 20 accepted indicators with the highest between-round agreement

uncorrected proof

This chart shows the percentage agreement in expert responses across the two Delphi rounds for the highest-ranked accepted indicators. Agreement ranged from 80% to 100%: eight indicators reached full agreement, seven reached 90%, and the remaining five shown reached 80%. These results indicate appropriate convergence in expert judgments and robustness of the resulting consensus. Because the Wilcoxon test was applied separately to each of the 29 retained indicators, the resulting p-values were adjusted for multiple comparisons using the Benjamini-Hochberg false discovery rate procedure. After adjustment, no indicator exhibited a statistically significant shift between rounds (smallest adjusted $p = 0.554$). The two smallest unadjusted values, obtained for IP15 ($p = 0.045$) and LG13 ($p = 0.047$), both correspond to an adjusted value of 0.554 and are reported for completeness only; both indicators met the three acceptance criteria and were retained in the final set. Details of the paired Wilcoxon test for the indicators are presented in Table 13.

Table 13. Paired Wilcoxon Test results for assessing indicator changes between the two Delphi rounds

Item	Median_Shift	IQR_Change	Agreement S_i (%)	Wilcoxon_P
LG1	0.000	-0.187	100	0.148
LG4	0.000	0.000	100	NA
IP7	0.000	-0.250	100	0.185
IP15	0.250	-0.250	100	0.045
F1	0.125	-0.250	100	0.269
C3	0.250	-0.250	100	0.130
F4	0.125	0.000	100	0.102
F6	0.250	0.145	100	0.198
IP10	0.000	-0.187	90	0.173
IP16	0.000	-0.250	90	0.088
LG13	0.125	-0.250	90	0.047
LG7	0.000	-0.187	90	0.853
LG8	0.250	0.062	90	0.222
F5	0.000	0.250	90	0.173
IP1	0.125	0.083	90	0.831
LG5	0.000	-0.250	80	0.409
IP6	0.000	-0.250	80	0.587
IP4	0.000	0.062	80	0.586
IP8	0.000	0.000	80	0.831
IP12	0.250	0.000	80	0.829

Uncorrected proof

F7	0.000	0.250	80	0.357
F9	0.000	0.000	80	0.831
C1	0.000	0.250	80	0.580
C2	0.000	0.000	80	0.831
LG2	0.250	0.145	80	0.524
C7	0.000	0.145	80	0.587
IP2	-0.166	0.145	80	0.276
F2	-0.041	0.083	80	0.798
F3	-0.041	0.083	80	0.441

NA = no variation between rounds; the Wilcoxon statistic is undefined.

Benjamini-Hochberg adjusted p-values are available from the authors; the smallest adjusted value is 0.554.

4.4. Final ranking of indicators based on CRI

After Min-Max normalization of the two components (C, DF) to the uniform interval [0, 1], the 29 accepted indicators were ranked according to the Composite Ranking Index ($CRI_i = 0.50 \times \hat{C}_i + 0.50 \times \hat{DF}_i$). Table 14 presents the CRI score and rank of each indicator by the four GBSC perspectives.

uncorrected proof

Table 14. Final ranking of indicators based on the composite ranking index (CRI)

Item	Mean	Median	Consensus	Defuzzified	Consensus_Scaled	Defuzzified_Scaled	Composite_Score (CRI)	Threshold ($\bar{d}$)	Final_Rank
LG4	1.000	1.000	100	0.890	1.000	1.000	1.000	0.046	1
LG1	1.000	1.000	100	0.890	1.000	1.000	1.000	0.084	2
IP7	0.966	1.000	100	0.867	1.000	0.800	0.900	0.090	3
IP15	0.966	1.000	100	0.867	1.000	0.800	0.900	0.116	4
F1	0.933	1.000	100	0.845	1.000	0.599	0.799	0.097	5

C3	0.933	1.000	100	0.845	1.000	0.599	0.799	0.110	6
IP16	1.000	1.000	90	0.890	0.500	1.000	0.750	0.120	7
IP10	1.000	1.000	90	0.890	0.500	1.000	0.750	0.132	8
LG13	1.000	1.000	90	0.890	0.500	1.000	0.750	0.133	9
F4	0.900	1.000	100	0.823	1.000	0.400	0.700	0.105	10
F6	0.866	1.000	100	0.800	1.000	0.200	0.600	0.116	11
LG7	0.933	1.000	90	0.845	0.500	0.599	0.549	0.111	12
F5	0.900	1.000	90	0.823	0.500	0.400	0.450	0.084	13
LG8	0.900	1.000	90	0.823	0.500	0.400	0.450	0.156	14
LG5	0.966	1.000	80	0.867	0.000	0.800	0.400	0.146	15
IP1	0.866	1.000	90	0.800	0.500	0.200	0.350	0.116	16
IP6	0.933	1.000	80	0.845	0.000	0.599	0.299	0.140	17
F7	0.900	1.000	80	0.823	0.000	0.400	0.200	0.117	18
C1	0.900	1.000	80	0.823	0.000	0.400	0.200	0.124	19
IP4	0.900	1.000	80	0.823	0.000	0.400	0.200	0.130	20
C2	0.900	1.000	80	0.823	0.000	0.400	0.200	0.143	21
F9	0.900	1.000	80	0.823	0.000	0.400	0.200	0.143	22
IP8	0.900	1.000	80	0.823	0.000	0.400	0.200	0.143	23
IP12	0.900	1.000	80	0.823	0.000	0.400	0.200	0.169	24
C7	0.866	1.000	80	0.800	0.000	0.200	0.100	0.131	25
LG2	0.866	1.000	80	0.800	0.000	0.200	0.100	0.165	26
IP2	0.833	0.833	80	0.778	0.000	0.000	0.000	0.124	27
F3	0.833	0.833	80	0.778	0.000	0.000	0.000	0.143	28
F2	0.833	0.833	80	0.778	0.000	0.000	0.000	0.150	29

uncorrected proof

Note: Where two or more indicators obtained the same Composite Ranking Index, ties were resolved by an explicit hierarchy applied uniformly across all tied bands: (i) the lower mean geometric distance d-bar, indicating closer agreement with the group judgment; (ii) the higher between-round agreement index S_i; and (iii) alphabetical order of the indicator code. The composition of the highest-ranked group is invariant to this rule; only the ordering within each tied band is affected.

The composite-ranking results in Table 14 show that LG4 and LG1, with scores of 1.000, rank first and second, followed by IP7 and IP15, with scores of 0.900. The six highest-ranked indicators were LG4, LG1, IP7, IP15, F1, and C3, with CRI scores ranging from 0.799 to 1.000. In terms of perspective distribution, internal processes (IP), with 10 indicators, accounted for the largest share of the 29 final indicators, followed by the

financial (F), learning and growth (LG), and customer (C) perspectives, indicating the breadth of operational coverage in the GBSC, although the two highest-ranked indicators belong to the learning and growth perspective.

By contrast, IP2, F2, and F3 received the lowest CRI scores, while LG2 and C7 also ranked near the bottom.

Table 15. Friedman test and Kendall's coefficient of concordance by perspective (Round-2 data, N = 10 experts)

Perspective	k	df	Chi-square	p	Kendall's W
Financial (F)	8	7	7.159	0.413	0.102
Customer (C)	4	3	4.318	0.229	0.144
Internal processes (IP)	10	9	2.618	0.978	0.029
Learning and growth (LG)	7	6	1.484	0.961	0.025

Note: The Friedman test was computed on the Round-2 responses of the 10 second-round experts, with k denoting the number of accepted indicators in each perspective and df = k - 1. Kendall's W, obtained as chi-square /(N x df), is reported as the effect size. None of the four tests reached significance at the 0.05 level, indicating that although the experts converged on which indicators should be retained, they did not converge on a common ordering of relative priority within each perspective. For this reason the mean ranks reported in Table 16 are presented for description only and without a pooled test statistic, and the final prioritization rests on the Composite Ranking Index.

Table 16. Descriptive mean ranks of the 29 final indicators across the second-round experts

Item	Mean_Rank
LG1	18.600
LG4	18.600
LG13	18.600
IP10	18.600
IP16	18.600
IP7	17.150
IP15	17.150
LG5	17.150
F1	15.700
C3	15.700
IP6	15.700
LG7	15.700
F4	14.250
F5	14.250
F7	14.250

Uncorrected proof

F9	14.250
C1	14.250
C2	14.250
IP4	14.250
IP8	14.250
IP12	14.250
LG8	14.250
F6	12.800
C7	12.800
IP1	12.800
LG2	12.800
IP2	11.350
F2	11.350
F3	11.350

In Table 16 a higher mean rank denotes a higher priority. The mean ranks were obtained by ranking the 29 indicators within each second-round expert's responses and averaging those ranks across the ten experts; they differ from the CRI ordering because the CRI also incorporates the consensus percentage. As reported in Table 15, none of the four within-perspective Friedman tests reached significance, and Kendall's W was low in every perspective; the experts therefore converged on which indicators to retain but not on a single ordering of relative priority. The mean ranks are consequently reported for description only and are not accompanied by a pooled test statistic, since the Friedman test was computed within each perspective, as explained in Section 3.5.5, and the final prioritization was based on the Composite Ranking Index.

4.5. Final GBSC conceptual framework

According to Table 17, the final Green Balanced Scorecard (GBSC) conceptual framework was developed based on 29 accepted indicators across four perspectives. The final structure comprises the Internal Processes perspective with 10 indicators, the Financial perspective with 8 indicators, the Customer perspective with 4 indicators, and the Learning & Growth perspective with 7 indicators. Drawing on five theoretical foundations (the BSC, the TBL and the NRBV as the main lenses, and stakeholder and legitimacy theory as complementary lenses) and based on expert consensus obtained through the Fuzzy Delphi Method, this framework provides broad coverage of the financial, operational, customer-oriented, and infrastructural dimensions of organizational environmental performance. The framework may provide a basis for

Uncorrected proof

developing tools for measurement, monitoring, and decision-making in the area of organizational environmental performance.

Table 17. The final GBSC conceptual framework by perspective and indicator

Code	Green Internal-Process Perspective - 10 Indicators	Code	Green Financial Perspective - 8 Indicators
IP1	Green manufacturing process (improving the environmental performance of production processes, adopting new and green technologies, reducing waste, and improving operational efficiency with a green approach)	**F1**	Green product and service revenue (eco-products and services and the proportion of sales of environmentally friendly products)
IP2	Sustainable energy management (consumption of renewable energy and energy efficiency)	**F2**	Green ancillary revenues (from recycling manufactured products and expanding revenue opportunities from environmental activities)
IP4	Recycled-material management (use of recycled materials in production and the recycling rate for production and office materials)	**F3**	Environmental non-compliance costs (pollution-related fines and costs resulting from failure to address environmental activities)
IP6	Reduction of production-generated wastewater	**F4**	Environmental operating costs (disposal costs and costs of environmental initiatives)
IP7	Reduction of greenhouse gas emissions	**F5**	Resource-consumption optimization (maximizing energy savings and cost

uncorrected proof

			savings from environmental improvement)
IP8	Waste management	**F6**	Waste-cost management (reducing operating costs by reducing waste and financial participation in reducing potential waste)
IP10	Use of a green supply chain (green supplier selection and relationships with sustainable suppliers)	**F7**	Investment in sustainable technologies (technologies effective in conserving natural resources such as water and energy)
IP12	Assessment of supplier environmental performance	**F9**	Return on green investment (environmental efficiency and return on environmental investment)
IP15	Recycling and extension of the raw-material cycle		
IP16	Product reuse rate		
Code	**Green Learning and Growth Perspective - 7 Indicators**	**Code**	**Green Customer Perspective - 4 Indicators**
LG1	Development of employees' knowledge of the organization's environmental impacts and environmental management procedures	C1	Provision of environmentally friendly products (environmentally safe products, provision of clean products to customers, and the value of eco-products)
LG2	Employees' environment-specific technical skills	C2	Sustainable product design and packaging (green product design and packaging)

uncorrected proof

LG4	Access to information and green knowledge sharing	C3	Green brand and corporate environmental image (presenting the image of an environmentally responsible company and the brand reputation of eco-products)
LG5	Technological innovations and applied research related to product planning, packaging, and production processes to solve environmental problems	C7	Green customer satisfaction (satisfaction and value creation for green customers)
LG7	Research and development for green technologies		
LG8	Hours allocated to research and development in environmental design		
LG13	Linking employee incentives to environmental objectives		

uncorrected proof

5. Discussion, conclusion, limitations, and recommendations

5.1. Overview

This section analyzes and interprets the findings of Section 4 in light of the theoretical foundations and previous studies. The results are first examined across the four GBSC perspectives, followed by a comparison of the final framework with previous studies, the theoretical and practical implications, limitations, research recommendations, and the overall conclusion.

5.2. Discussion of findings by the four GBSC perspectives

5.2.1. Financial perspective

According to Table 14 (final CRI ranking), 'green ancillary revenues' (F2) and 'environmental non-compliance costs' (F3) received the lowest ranks in the financial

perspective, while 'investment in sustainable technologies' (F7) and 'return on green investment' (F9) also ranked in the lower half. In addition, F10 (increased profitability through a green product-differentiation strategy) was removed in the first Delphi round (Table 11).

The study's findings regarding the importance of green investment are consistent with previous studies. Chen and Ma (2021) showed that green investment is positively related to the financial performance of energy companies and that this effect is strengthened through improved environmental performance and fewer violations. Rahat and Nguyen (2022) also stated that green investments in emerging markets can generate more favorable risk-adjusted returns. In addition, based on an examination of polluting firms listed on China's stock exchanges, Xing et al. (2025) showed that environmental violations are associated with increased financial distress risk through financing constraints and weak internal controls.

Despite this consistency, the present findings differ from some empirical studies. While Xing et al. (2025) regard environmental non-compliance costs as a factor affecting financial risk, F3 received a low score in the CRI ranking. This result does not indicate that the indicator is unimportant; rather, it indicates that, compared with the other financial indicators, this criterion received lower consensus and priority among the experts. In addition, Shevchenko (2021) showed that environmental penalties do not always lead to improved environmental performance and may have complex financial and managerial consequences.

The removal of F10 also indicates that experts regard green profitability as meaningful when it is linked to specific environmental and operational indicators. This view is consistent with Chen and Ma (2021), who regard the financial benefits of green actions primarily as a long-term outcome of improved environmental performance and reduced organizational violations.

5.2.2. Customer perspective

In the final GBSC framework, four indicators-"green brand" (C3), "provision of environmentally friendly products" (C1), "sustainable design and packaging" (C2), and "green customer satisfaction" (C7)-were retained in the customer perspective. By contrast, C5 (environmental risk management and legal compliance) and C6 (strength of regulatory relationships) were removed in the first Delphi round (Table 11), and C4 (enhanced media coverage) in the second (Table 12). This pattern indicates that the

uncorrected proof

experts prioritized indicators related to direct customer perception, product experience, and brand value over regulatory and precautionary indicators in the customer perspective. The confirmed importance of the green brand and sustainable products is consistent with previous studies. Khan et al. (2022b) showed that green practices increase green brand equity by strengthening brand attachment, whereas green skepticism can weaken this relationship. Ha (2022) also stated that greenwashing indirectly affects brand equity through its effects on brand image and green trust, which is consistent with the argument advanced in Section 2.3.2 that sustainability claims alone do not guarantee market success. In addition, the confirmation of the sustainable product and packaging indicators is consistent with the findings of Pogačar and Gregor-Svetec (2025) concerning the role of eco-design, reduced material consumption, and the use of recyclable materials in packaging.

Despite the importance of regulatory indicators in some studies, the present results show a different pattern of prioritization. Du et al. (2024) emphasize the role of regulatory frameworks in strengthening the effect of ESG performance on risk management, and Andreoli et al. (2025) likewise address the importance of strategic accountability and ESG-related litigation risk; however, indicators C4, C5 and C6 did not reach the level of consensus required to enter the final framework in the Delphi process. This reflects the experts' emphasis on observable and customer-oriented dimensions of green performance in designing the GBSC customer perspective (see also Table 1).

Uncorrected proof

5.2.3. Internal-process perspective

According to Table 14, the internal-process perspective, with 10 indicators, accounts for the largest share of the final GBSC framework. IP7 (reduction of greenhouse gas emissions) and IP15 (recycling and extension of the raw-material cycle) received the highest CRI scores within this perspective and ranked third and fourth overall, followed, within this perspective, by IP16 (product reuse rate) and IP10 (use of a green supply chain). The other confirmed indicators were the green manufacturing process (IP1), reduction of production-generated wastewater (IP6), recycled-material management (IP4), waste management (IP8), assessment of supplier environmental performance (IP12), and sustainable energy management (IP2). This concentration reflects the breadth of operational coverage in the GBSC.

The importance of these indicators is consistent with the literature and international standards. Panagiotopoulou et al. (2022) showed that energy and raw-material

consumption are among the principal contributors to carbon footprints in manufacturing systems. Voukkali et al. (2023) also identified waste management as a key component of the circular economy. In addition, ISO 50001, GRI 305, and GRI 306 emphasize the importance of energy management, greenhouse gas emissions, and waste in environmental performance assessment. In the supply-chain context, Abuzaid et al. (2024) and Li and Zhong (2024) have also emphasized the role of green supplier selection and supply-chain integration in improving environmental performance.

Despite this consistency, differences in prioritization were observed. Although ISO 50001 and Panagiotopoulou et al. (2022) emphasize the importance of energy management, IP2 received a low score in the CRI ranking. Similarly, despite their theoretical importance in the circular economy, IP4 and IP8 were assigned lower priority than indicators related to the direct reduction of emissions and material recycling. This result indicates that experts prioritized indicators with more direct effects on environmental performance over some managerial and supporting indicators.

Finally, IP11 (use of environmental auditing and assessment of suppliers) and IP13 (environmentally friendly packaging) were removed in the first Delphi round (Table 11). Their removal indicates that, compared with the other internal-process indicators, these criteria did not receive sufficient consensus and priority for inclusion in the final GBSC framework.

uncorrected proof

5.2.4. Learning and growth perspective

Within the learning and growth perspective, seven indicators were retained in the final GBSC framework: development of employees' knowledge of the organization's environmental impacts and environmental management procedures (LG1), employees' environment-specific technical skills (LG2), access to information and green knowledge sharing (LG4), technological innovations and applied research (LG5), research and development for green technologies (LG7), hours allocated to research and development in environmental design (LG8), and linking employee incentives to environmental objectives (LG13). This result indicates that experts regard green human capital, technological innovation capacity, knowledge sharing, and the alignment of incentives with environmental objectives as supporting factors for achieving organizational environmental objectives.

This finding is consistent with Dzhengiz and Niesten (2020), who emphasize the role of employees' environmental competencies, organizational absorptive capacity, and

organizational learning in developing environmental performance. Bresciani et al. (2023) also showed that the use of Environmental Management Accounting (EMA), environmental knowledge management, top-management support, and the creation of a green work environment contribute to improved organizational environmental performance. Accordingly, the findings of the present study emphasize the role of human capital, organizational knowledge, and knowledge sharing as drivers of green learning and long-term sustainability within the GBSC.

5.3. Comparison with previous studies

Compared with previous versions of the GBSC, the framework developed in this study has three key distinctions:

First: simultaneous integration of international standards (ISO 14001, ISO 50001, ISO 14051 and the GRI standards) with the recent scientific literature, whereas previous studies such as Moraga et al. (2022) largely focused on one of these two routes.

Second: coverage of the interdisciplinary dimension through a three-disciplinary expert panel (accounting, management, and agricultural economics), consistent with the interdisciplinary nature of the GBSC.

Third: use of the Composite Ranking Index (CRI), which simultaneously aggregates the two dimensions of importance and consensus into a single crisp value and prevents the ranking from becoming one-dimensional. This methodological innovation has been less frequently observed in previous Fuzzy Delphi studies.

5.4. Theoretical contribution of the study

Extension of the green BSC literature by presenting a four-perspective framework with 29 accepted indicators supported by expert consensus and international standards.

Strengthening the theoretical bridge among green accounting, sustainability management, and environmental economics through an interdisciplinary panel.

Introduction of the CRI as a new methodological tool in consensus-based research that can be generalized to other areas of multi-criteria decision-making.

Explicit mapping of the five theoretical foundations onto the four perspectives and their indicators, with a stated assignment criterion for each foundation (Table 1).

Operationalization of the financial perspective in terms of recognized accounting treatments and disclosure frameworks, thereby linking the GBSC to environmental management accounting and to material flow cost accounting (Table 3).

5.5. Practical contribution of the study

Uncorrected proof

Senior managers: the proposed framework can serve as a conceptual basis for designing tools to measure financial and environmental performance simultaneously.

accounting and sustainability-reporting units: as specified in Table 3, the financial indicators map onto recognized accounting treatments (IAS 37, IAS 16, IFRS 15) and onto GRI and ISSB disclosure requirements, while the internal-process indicators supply the physical quantity data required for material flow cost accounting under ISO 14051.

policy-makers: Identifying the key indicators in each perspective facilitates the development of incentive regulations for green companies and the design of environmental scoring systems.

human resources and training units: Indicators in the learning and growth perspective provide a clear roadmap for designing green human-capital development programs.

6. Research limitations

Reduction in the number of experts in the second round: Although the number of experts in both rounds (15 in the first round and 10 in the second) falls within the standard range, the decline in participation from the first to the second round somewhat limited the generalizability of the findings, although the non-response analysis reported in Appendix D indicates that it did not materially bias the indicator set; lack of time was the reason reported by all five non-respondents, which is the most frequent reason for discontinuation in Delphi studies (Niederberger et al., 2025).

Selection of experts from within the country: Selecting experts exclusively from the domestic academic community did not allow the full representation of international perspectives or practical industry experience.

Subjective nature of the Fuzzy Delphi Method: Although fuzzy logic reduces ambiguity in judgments, the Delphi method remains rooted in experts' subjective judgment and cannot be made entirely free from human bias.

Weak concordance in the relative ranking of indicators: Although expert consensus on the retention of indicators was high, the within-perspective Friedman tests reported in Table 15 were not significant and Kendall's W was low in every perspective. The prioritization presented here should therefore be regarded as indicative rather than definitive and merits re-examination with a larger panel.

Lack of empirical testing: The proposed framework was validated solely on the basis of expert consensus and has not been empirically assessed in real companies.

Absence of a formal inter-coder reliability coefficient: The content analysis through which the 76 extracted components were reduced to 47 indicators was conducted by a

uncorrected proof

single coder and reviewed by the supervising author, without the computation of an inter-coder reliability coefficient. The resulting indicators were, however, subsequently validated by an independent panel of experts through two rounds of Fuzzy Delphi screening.

Timing of the survey: The expert panel was surveyed in autumn 2024; although the literature was updated during the revision process, the consensus judgments reported here reflect expert views at that point in time.

7. Recommendations for future research

Test the cause-and-effect relationships between the four perspectives in manufacturing and service companies using partial least squares structural equation modeling (PLS-SEM) with organizational data.

Expand the expert panel to specialists from different countries and compare the findings across different cultural and regional contexts.

Integrate the framework with multi-criteria decision-making methods such as fuzzy ANP, fuzzy DEMATEL, or fuzzy TOPSIS to examine interdependencies among the four perspectives.

Conduct a longitudinal study over a 3- to 5-year period to evaluate the effects of implementing the framework on organizations' financial and environmental performance.

Align the framework with emerging sustainability-reporting standards such as ISSB and CSRD to accommodate global developments in non-financial reporting.

Examine the role of variables such as organizational size, industry type, ownership structure, and level of environmental maturity in the effectiveness of this framework.

uncorrected proof

8. Overall conclusion

By integrating targeted document analysis with a two-stage Fuzzy Delphi Method, the present study developed a four-perspective framework for the Green Balanced Scorecard (GBSC), in which 29 key indicators-the internal-process (IP) perspective with 10 indicators, the financial (F) perspective with 8 indicators, the customer (C) perspective with 4 indicators, and the learning and growth (LG) perspective with 7 indicators-were accepted based on expert consensus, vertex-based geometric distance, and defuzzification scores. The indicator-screening path from 47 initial indicators to 39 indicators in the first round and ultimately to 29 accepted indicators in the second round reflects a systematic and transparent screening process.

From a methodological perspective, the combination of centroid defuzzification, vertex-based geometric distance, and the CRI as an integrated analytical package simultaneously reflected the two dimensions of importance and consensus in the indicator-prioritization process. Overall, the proposed framework may serve as a conceptual model and a basis for developing interdisciplinary decision-making tools, bridge the literature on green accounting, sustainability management, and environmental economics, and guide future empirical research in this field.

Accordingly, the framework is presented as the output of the first stage of a research program: the strategic cause-and-effect chains linking the four perspectives were not modeled or tested in this study and constitute the next stage of the research.

Ethics statement

This study involved no clinical, medical or vulnerable participants and no personal or sensitive data. It was limited to the professional opinions of academic experts, elicited anonymously. Under the applicable national regulations, anonymous expert surveys of this kind do not require formal institutional ethics approval; the study nevertheless adhered to the principles of the Declaration of Helsinki.

Participation was entirely voluntary. The cover letter accompanying the Round-1 questionnaire, reproduced in Appendix A, informed each expert of the purpose and academic context of the study and invited their voluntary professional judgment; experts gave their consent by returning the completed questionnaire. No personally identifiable information was recorded at any stage, and experts are identified in the analysis only by anonymized codes (E1 to E15). All results are reported in aggregate form.

uncorrected proof

References

Abuzaid, A. N., Alateeq, M. M., Madadha, S. A. M., Al Sharari, F. E. A. ,& Alsbou, M. K. (2024). The effect of suppliers' green and traditional selection criteria in supply chain management on purchasing firms' performance. *Sustainability*, 16(15), Article 6276. https://doi.org/10.3390/su16156276

Al Hanini, E. A. (2021). Does the environmental internal audit impact the achieving of sustainable development in industrial companies listed on the Amman stock exchange?. *International Journal of Financial Research*, 12(3), 220-229. https://doi.org/10.5430/ijfr.v12n3p220

Alghababsheh, M., Butt, A. S. , & Moktadir, M. A. (2022). Business strategy, green supply chain management practices, and financial performance: A nuanced empirical examination. *Journal of Cleaner Production,* 380, Article 134865. https://doi.org/10.1016/j.jclepro.2022.134865

Aljamal, D., Salem, A., Khanna, N. , & Hegab, H. (2024). Towards sustainable manufacturing: A comprehensive analysis of circular economy key performance indicators in the manufacturing industry. *Sustainable Materials and Technologies*, 40, Article e00953. https://doi.org/10.1016/j.susmat.2024.e00953

Amoako, G. K., Bawuah, J., Asafo-Adjei, E. , & Ayimbire, C. (2023). Internal audit functions and sustainability audits: Insights from manufacturing firms. *Cogent Business*

Uncorrected proof

& *Management,* 10(1), Article 2192313. https://doi.org/10.1080/23311975.2023.2192313

Andreoli, C., Cremasco, C., Falivena, C. , & Brunelli, S. (2025). ESG and impact litigation: identifying and governing the causes through strategic accountability patterns. *Management Decision,* 63(2), 640-664. https://doi.org/10.1108/MD-10-2023-2008

Bendig, D., Kleine-Stegemann, L. , & Gisa, K. (2023). The green manufacturing framework—A systematic literature review. *Cleaner Engineering and Technology*, 13, Article 100613. https://doi.org/10.1016/j.clet.2023.100613

Betts, K., Gutierrez-Franco, E. , & Ponce-Cueto, E. (2022). Key metrics to measure the performance and impact of reusable packaging in circular supply chains. *Frontiers in Sustainability,* 3, Article 910215. https://doi.org/10.3389/frsus.2022.910215

Bouzon, M., Govindan, K., Rodriguez, C. M. T. , & Campos, L. M. (2016). Identification and analysis of reverse logistics barriers using fuzzy Delphi method and AHP. *Resources, Conservation and Recycling,* 108, 182-197. https://doi.org/10.1016/j.resconrec.2015.05.021

Bresciani, S., Rehman, S. U., Giovando, G. , & Alam, G. M. (2023). The role of environmental management accounting and environmental knowledge management practices influence on environmental performance: mediated-moderated model. *Journal of Knowledge Management,* 27(4), 896-918. https://doi.org/10.1108/JKM-12-2021-0953

Camañes, V., Tobajas, R. , & Fernandez, A. (2024). Methodology of eco-design and software development for sustainable product design. *Sustainability*, 16(7), 2626. https://doi.org/10.3390/su16072626

Cao, J., Huang, Z. , & Kristanto, A. B. (2026). Green hunter or green detractor? Strategic orientations and green revenues. *British Journal of Management*, 37(3), Article e70063. https://doi.org/10.1111/1467-8551.70063

Chai, S., Cao, M., Li, Q., Ji, Q. , & Liu, Z. (2023). Exploring the nexus between ESG disclosure and corporate sustainable growth: Moderating role of media attention. *Finance Research Letters,* 58, Article 104519. https://doi.org/10.1016/j.frl.2023.104519

Chang, A. S. , & Tsai, C. Y. (2015). Sustainable design indicators: Roadway project as an example. *Ecological Indicators*, 53, 137-143. https://doi.org/10.1016/j.ecolind.2015.01.036

Chen, R., He, X. , & Shirani Bidabadi, F. (2023). Corporate environmental compliance in china: from social responsibility to soft law. *Sustainability*, 15(3), Article 2379. https://doi.org/10.3390/su15032379

uncorrected proof

Chen, S. J., & Hwang, C. L. (1992). Fuzzy multiple attribute decision making methods. In Fuzzy multiple attribute decision making: Methods and applications (pp. 289–486). Springer. https://doi.org/10.1007/978-3-642-46768-4

Chen, S., Jiang, W., Li, X. , & Gao, H. (2021). Effect of employees' perceived green HRM on their workplace green behaviors in oil and mining industries: Based on cognitive-affective system theory. *International Journal of Environmental Research and Public Health,* 18(8), Article 4056. https://doi.org/10.3390/ijerph18084056

Chen, Y. , & Ma, Y. (2021). Does green investment improve energy firm performance?. *Energy Policy*, 153, Article 112252. https://doi.org/10.1016/j.enpol.2021.112252

Cheng, C. H. , & Lin, Y. (2002). Evaluating the best main battle tank using fuzzy decision theory with linguistic criteria evaluation. *European Journal of Operational Research*, 142(1), 174-186. https://doi.org/10.1016/S0377-2217(01)00280-6

Cheng, C. H. , & Mon, D. L. (1994). Evaluating weapon system by analytical hierarchy process based on fuzzy scales. *Fuzzy Sets and Systems*, 63(1), 1-10. https://doi.org/10.1016/0165-0114(94)90140-6

Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, 6(3), 399-403. https://doi.org/10.3844/jssp.2010.399.403

Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. Performance Improvement, 56(10), 15-23. https://doi.org/10.1002/pfi.21727

Cristea, M., Noja, G. G., Drăcea, R. M., Iacobuță-Mihăiță, A. O. , & Dorożyński, T. (2024). ESG investment strategies and the financial performance of European agricultural companies: a new modelling approach. *Journal of Business Economics and Management*, 25(6), 1283-1307. https://doi.org/10.3846/jbem.2024.23010

D'Almeida, A. P. , & de Albuquerque, T. L. (2024). Innovations in food packaging: from bio-based materials to smart packaging systems. *Processes*, 12(10), Article 2085. https://doi.org/10.3390/pr12102085

Dajani, J. S., Sincoff, M. Z. , & Talley, W. K. (1979). Stability and agreement criteria for the termination of Delphi studies. *Technological Forecasting and Social Change*, 13(1), 83-90. https://doi.org/10.1016/0040-1625(79)90007-6

De Kruijf, B. J., Edzes, A., van Dijk, J. , & Waslander, S. (2025). Learning climate and financial incentives: A study on training participation in Dutch organizations. *Journal of Adult and Continuing Education*, 31(3), 734-751. https://doi.org/10.1177/14779714241278108

uncorrected proof

Delaney, E., Liu, W., Zhu, Z., Xu, Y. , & Dai, J. S. (2022). The investigation of environmental sustainability within product design: a critical review. *Design Science*, 8, Article e15. https://doi.org/10.1017/dsj.2022.11

Dhanorkar, S. , & Siemsen, E. (2021). How nudges lead to improved energy efficiency in manufacturing: Evidence from archival data and a field study. *Production and Operations Management,* 30(10), 3735-3757. https://doi.org/10.1111/poms.13479

Dierkes, S. , & Siepelmeyer, D. (2025). Material flow cost accounting with multiple inefficiency factors and recycling. *Schmalenbach Journal of Business Research*, 77(1), 57-93. https://doi.org/10.1007/s41471-024-00197-z

Du, Q., Sun, Z., Goodell, J. W., Du, A. M. , & Yang, T. (2024). Ecological risk and corporate sustainability: Examining ESG performance, risk management, and productivity. *International Review of Financial Analysis*, 96, Article 103551. https://doi.org/10.1016/j.irfa.2024.103551

Dzhengiz, T. , & Niesten, E. (2020). Competences for environmental sustainability: A systematic review on the impact of absorptive capacity and capabilities. *Journal of Business Ethics*, 162(4), 881-906. https://doi.org/10.1007/s10551-019-04360-z

Earnhart, D. , & Harrington, D. R. (2021). Effects of audit frequency, audit quality, and facility age on environmental compliance. *Applied Economics*, 53(28), 3234-3252. https://doi.org/10.1080/00036846.2020.1854449

Elkington, J. (1997). Cannibals with forks: The triple bottom line of 21st century business. Capstone.

Farooque, M., Zhang, A., Thürer, M., Qu, T. , & Huisingh, D. (2019). Circular supply chain management: A definition and structured literature review. *Journal of Cleaner Production,* 228, 882-900. https://doi.org/10.1016/j.jclepro.2019.04.303

França, A., López-Manuel, L., Sartal, A. , & Vázquez, X. H. (2023). Adapting corporations to climate change: How decarbonization impacts the business strategy–performance nexus. *Business Strategy and the Environment*, 32(8), 5615-5632. https://doi.org/10.1002/bse.3439

Freeman, R. E. (1984). Strategic management: A stakeholder approach. Pitman.

García-Quevedo, J. ,& Jové-Llopis, E. (2021). Environmental policies and energy efficiency investments. An industry-level analysis. *Energy Policy*, 156, Article 112461. https://doi.org/10.1016/j.enpol.2021.112461

Global Reporting Initiative. (GRI). (2021). GRI 2: General disclosures 2021. USA /Canada. https://www.globalreporting.org/standards/

Uncorrected proof

Global Reporting Initiative. (GRI). (2020). *GRI 306: Waste 2020*. USA /Canada. https://www.globalreporting.org/standards/

Global Reporting Initiative. (GRI). (2021). *GRI 303: Water and effluents 2018*. USA /Canada. https://www.globalreporting.org/standards/

Global Reporting Initiative. (GRI). (2021). *GRI 305: Emissions 2016*. USA /Canada. https://www.globalreporting.org/standards/

Gull, A. A., Atif, M., Ahsan, T. , & Derouiche, I. (2022). Does waste management affect firm performance? International evidence. *Economic Modelling*, 114, Article 105932. https://doi.org/10.1016/j.econmod.2022.105932

Gusmerotti, N. M., Todaro, N. M., Tosi, D. , & Testa, F. (2023). Green work climate, work meaningfulness and supervisor environmental priority: A social exchange perspective on employees' eco-initiatives. *Journal of Cleaner Production*, 415, Article 137889. https://doi.org/10.1016/j.jclepro.2023.137889

Ha, M. T. (2022). Greenwash and green brand equity: The mediating role of green brand image, green satisfaction, and green trust, and the moderating role of green concern. *PloS One,* 17(11), Article e0277421. https://doi.org/10.1371/journal.pone.0277421

Hart, S. L. (1995). A natural-resource-based view of the firm. Academy of Management Review, 20(4), 986–1014. https://doi.org/10.5465/amr.1995.9512280033

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. Journal of Advanced Nursing, 32(4), 1008-1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x

Hidayat, I., Abbas, D. S., Lam, N. T. , & Sari, P. A. (2024). The role of environmental management accounting in mediating green innovation to firm value: Moderated by quality management. *International Journal of Energy Economics and Policy*, 14(3), 281-287. https://doi.org/10.32479/ijeep.15869

Hsu, T. H. , & Yang, T. H. (2000). Application of fuzzy analytic hierarchy process in the selection of advertising media. *Journal of Management and Systems,* 7(1), 19–39.

International Organization for Standardization. (ISO). (2015). *ISO 14001:2015 environmental management systems—Requirements with guidance for use*. ISO. Geneva, Switzerland. https://www.iso.org/standard/60857.html

International Organization for Standardization. (ISO). (2018). *ISO 50001:2018 energy management systems—Requirements with guidance for use*. ISO. Geneva, Switzerland

Unconnected proof

International Accounting Standards Board. (IASB). (2003). IAS 2: Inventories. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/list-of-standards/ias-2-inventories/

International Accounting Standards Board. (IASB). (2003). IAS 16: Property, plant and equipment. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/list-of-standards/ias-16-property-plant-and-equipment/

International Accounting Standards Board. (IASB). (1998). IAS 37: Provisions, contingent liabilities and contingent assets. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/list-of-standards/ias-37-provisions-contingent-liabilities-and-contingent-assets/

International Accounting Standards Board. (IASB). (1998). IAS 38: Intangible assets. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/list-of-standards/ias-38-intangible-assets/

International Accounting Standards Board. (IASB). (2014). IFRS 15: Revenue from contracts with customers. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/list-of-standards/ifrs-15-revenue-from-contracts-with-customers/

International Sustainability Standards Board. (ISSB). (2023). IFRS S2: Climate-related disclosures. IFRS Foundation. London, United Kingdom. https://www.ifrs.org/issued-standards/ifrs-sustainability-standards-navigator/ifrs-s2-climate-related-disclosures/

International Organization for Standardization. (ISO). (2011). ISO 14051:2011 Environmental management-Material flow cost accounting-General framework. ISO. Geneva, Switzerland. https://www.iso.org/standard/50986.html

Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R. , & Mieno, H. (1993). The max-min Delphi method and fuzzy Delphi method via fuzzy integration. *Fuzzy Sets and Systems,* 55(3), 241-253. https://doi.org/10.1016/0165-0114(93)90251-C

Jassem, S., Zakaria, Z. , & Che Azmi, A. (2022). Sustainability balanced scorecard architecture and environmental performance outcomes: A systematic review. *International Journal of Productivity and Performance Management*, 71(5), 1728-1760. https://doi.org/10.1108/IJPPM-12-2019-0582

Jezzini, Y., Assaad, R. H., Boufadel, M. , & Nassif, H. (2024). Assessing the costs and benefits of green infrastructure plans using agent-based modeling and scenario analysis: Evaluating social and economic values. *Journal of Urban Planning and Development*, 150(4), Article 04024050. https://doi.org/10.1061/JUPDDM.UPENG-4664

uncorrected proof

Kaplan, R. S. , & Norton, D. P. (1992). The balanced scorecard—Measures that drive performance. *Harvard Business Review,* 70(1), 71–79.

Karim, E., Elmarzouky, M. , & Shohaieb, D. (2025). Green revenue generation and sales contribution: Are consumers willing to bear the cost of sustainability?. *Business Strategy and the Environment*, 34(8), 10977-10996. https://doi.org/10.1002/bse.70146

Khalili, S., Kumar, P. , & Jones, L. (2024). Evaluating the benefits of urban green infrastructure: Methods, indicators, and gaps. *Heliyon*, 10(19), Article e38446. https://doi.org/10.1016/j.heliyon.2024.e38446

Khan, A. N., Mehmood, K. , & Kwan, H. K. (2024). Green knowledge management: A key driver of green technology innovation and sustainable performance in the construction organizations. *Journal of Innovation & Knowledge*, 9(1), Article 100455. https://doi.org/10.1016/j.jik.2023.100455

Khan, M. T., Idrees, M. D., Rauf, M., Sami, A., Ansari, A. , & Jamil, A. (2022a). Green supply chain management practices' impact on operational performance with the mediation of technological innovation. *Sustainability*, 14(6), Article 3362. https://doi.org/10.3390/su14063362

Khan, S. A. R., Sheikh, A. A., Ashraf, M. , & Yu, Z. (2022b). Improving consumer-based green brand equity: The role of healthy green practices, green brand attachment, and green skepticism. *Sustainability*, 14(19), Article 11829. https://doi.org/10.3390/su141911829

Klausmann, J., Krueger, P. , & Matos, P. (2024). The green transition: Evidence from corporate green revenues. *FEB-RN Research Paper*, (29), 25-40. https://doi.org/10.2139/ssrn.4850449

Kosta, B. , & Ramadani, V. (2025). The impact of circular economy practices on firm performance: Evidence from the Western Balkans. *Journal of Enterprising Communities: People and Places in the Global Economy,* 19(4), 966–990. https://doi.org/10.1108/JEC-12-2024-0271

Kuo, Y. K., Khan, T. I., Islam, S. U., Abdullah, F. Z., Pradana, M. , & Kaewsaeng-On, R. (2022). Impact of green HRM practices on environmental performance: The mediating role of green innovation. *Frontiers in Psychology*, 13, Article 916723. https://doi.org/10.3389/fpsyg.2022.916723

Lee, H. , & Wu, X. (2023). Green supplier selection and environmental performance of firms in the Chinese manufacturing industry: the roles of behavior and outcome controls. *Journal of Manufacturing Technology Management,* 34(7), 1141-1161. https://doi.org/10.1108/JMTM-12-2022-0425

Uncorrected proof

Li, J. , & Zhong, D. (2024). Comparing the impact of green supplier selection and integration on environmental performance: An analysis of the moderating role of government support. *Sustainability*, 16(16), 7228. https://doi.org/10.3390/su16167228

Li, R. , & Li, H. (2024). The impact of food packaging design on users' perception of green awareness. *Sustainability*, 16(18), 8205. https://doi.org/10.3390/su16188205

Ma, Y. , & Wang, J. (2024). Employee green innovation behavior based on the ability-motivation-opportunity framework: Fuzzy-set qualitative comparative analysis. *Frontiers in Sustainability,* 5, Article 1415832. https://doi.org/10.3389/frsus.2024.1415832

Martín-Gómez, A. M., Pineda-Ganfornina, M., Ávila-Gutiérrez, M. J., Agote-Garrido, A. , & Lama-Ruiz, J. R. (2024). Balanced scorecard for circular economy: A methodology for sustainable organizational transformation. *Sustainability*, 16(4), Article 1464. https://doi.org/10.3390/su16041464

Mastria, S., Vezzil, A. , & De Cesarei, A. (2024). Switching to reuse: The impact of information on consumers' choices for reusable food packaging. *Sustainability*, 16(14), Article 5937. https://doi.org/10.3390/su16145937

Mat Nawi, M. N., Fauzi, M. A., Ting, I. W. K., Wider, W. , & Amaka, G. B. (2025). Green information technology and green information systems: science mapping of present and future trends. *Kybernetes*, 54(6), 3136-3155. https://doi.org/10.1108/K-10-2023-2139

Moraga, G., Huysveld, S., De Meester, S. , & Dewulf, J. (2022). Resource efficiency indicators to assess circular economy strategies: A case study on four materials in laptops. *Resources, Conservation and Recycling,* 178, Article 106099. https://doi.org/10.1016/j.resconrec.2021.106099

Morea, D., Gattermann Perin, M., Kolling, C., De Medeiros, J. F. , & Duarte Ribeiro, J. L. (2023). Environmental product innovation and perceived brand value: The mediating role of ethical-related aspects. *Sustainability*, 15(14), Article 10996. https://doi.org/10.3390/su151410996

Moussa, A. S. , & Elmarzouky, M. (2023). Does capital expenditure matter for ESG disclosure? A UK perspective. *Journal of Risk and Financial Management,* 16(10), 429. https://doi.org/10.3390/jrfm16100429

Murray, T. J., Pipino, L. L. , & Van Gigch, J. P. (1985). A pilot study of fuzzy set modification of Delphi. *Human Systems Management*, 5(1), 76-80. https://doi.org/10.3233/HSM-1985-5111

uncorrected proof

Nadler, J. T., Weston, R., & Voyles, E. C. (2015). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. The Journal of General Psychology, 142(2), 71-89 .https://doi.org/10.1080/00221309.2014.994590

Nazir, Y. , & Tennojiya, T. (2026). Environmental conservation costs and operational efficiency: Evidence from Japanese manufacturing firms. *Okayama Economic Review,* 57(3), 93–109. https://doi.org/10.1787/583CAB4C-EN

Niederberger, M., Dieudonné, J., Schifano, J., & Vogt, S .(2025) .Participants' experiences with the Delphi method: An online evaluation of a three-round Delphi across German-speaking countries. BMJ Open, 15(12), e104073 https://doi.org/10.1136/bmjopen-2025-105878

Okoli, C. , & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, 42(1), 15-29. https://doi.org/10.1016/j.im.2003.11.002

Orazalin, N., Ntim, C. , & Narbaev, T. (2026). Waste management, green initiatives, and financial distress in heavily regulated environmental contexts: evidence from the United Kingdom. *Journal of Accounting & Organizational Change*, 22(1), 1-22. https://doi.org/10.1108/JAOC-06-2024-0200

Pajuelo-Moreno, M. L., Barroso-Méndez, M. J. , & Gallardo-Vázquez, D. (2024). Relationship between sustainability disclosure and corporate reputation: evidence from a meta-analysis. *Business Strategy and the Environment*, 33(8), 8593-8615. https://doi.org/10.1002/bse.3933

Panagiotopoulou, V. C., Stavropoulos, P. , & Chryssolouris, G. (2022). A critical review on the environmental impact of manufacturing: a holistic perspective. *The International Journal of Advanced Manufacturing Technology*, 118(1), 603-625. https://doi.org/10.1007/s00170-021-07980-w

Pogačar, A. , & Gregor-Svetec, D. (2025). Eco-friendly design for sustainable gift packaging. *Applied Sciences*, 15(6), Article 2973. https://doi.org/10.3390/app15062973

Postacchini, L., Cicconi, P., Ciarapica, F. E., Germani, M. , & Bevilacqua, M. (2021). A design method for improving assembly and environmental sustainability in packaging solutions: A case study in household appliances. *International Journal of Sustainable Engineering*, 14(4), 574-589. https://doi.org/10.1080/19397038.2021.1920644

Rahat, B. , & Nguyen, P. (2022). Risk-adjusted investment performance of green and black portfolios and impact of toxic divestments in emerging markets. *Energy Economics*, 116, Article 106423. https://doi.org/10.1016/j.eneco.2022.106423

uncorrected proof

Ren, Y. , & Li, B. (2022). Digital transformation, green technology innovation and enterprise financial performance: Empirical evidence from the textual analysis of the annual reports of listed renewable energy enterprises in China. *Sustainability*, 15(1), 712. https://doi.org/10.3390/su15010712

Saeed, S. , & Cek, K. (2025). Integrating corporate sustainability reporting and environmental management systems into environmental performance management: an audit perspective. *Applied Economics*, 57(58), 10292-10307. https://doi.org/10.1080/00036846.2024.2427424

Saeedi, M. , & Ashraf, B. N. (2024). The role of technology in promoting green finance: A systematic literature survey and the development of a framework. *Journal of Risk and Financial Management*, 17(10), 472. https://doi.org/10.3390/jrfm17100472

Sallnäs, U. , & Björklund, M. (2023). Green e-commerce distribution alternatives–a mission impossible for retailers?. *The International Journal of Logistics Management,* 34(7), 50-74. https://doi.org/10.1108/IJLM-07-2022-0271

Sarjana, S. (2023). Green transportation: Development opportunities in support of sustainable transportation. E3S Web of Conferences, 429, Article 03003. https://doi.org/10.1051/e3sconf/202342903003

Sendawula, K., Turyakira, P., Akileng, G. , & Vincent, B. (2024). Environmental knowledge, regulatory compliance mechanisms, and environmental sustainability practices of manufacturing small and medium enterprises (SMEs) in Uganda. *Journal of Innovation and Entrepreneurship*, 13(1), 75. https://doi.org/10.1186/s13731-024-00426-z

Sherif, Z., Sarfraz, S., Jolly, M. , & Salonitis, K. (2022). Identification of the right environmental KPIs for manufacturing operations: Towards a continuous sustainability framework. *Materials*, 15(21), 7690. https://doi.org/10.3390/ma15217690

Shevchenko, A. (2021). Do financial penalties for environmental violations facilitate improvements in corporate environmental performance? An empirical investigation. *Business Strategy and the Environment*, 30(4), 1723-1734. https://doi.org/10.1002/bse.2711

Shi, T., Han, G., Ma, X., Mao, H., Chen, C., Han, Z., ... , & Gong, W. (2023). Quantifying factory-scale CO2/CH4 emission based on mobile measurements and EMISSION-PARTITION model: Cases in China. *Environmental Research Letters*, 18(3), Article 034028. https://doi.org/10.1088/1748-9326/acbce7

uncorrected proof

Slunge, D., Miguel, M., Lindahl, L. , & Backhaus, T. (2023). The implementation of the substitution principle in European chemical legislation: a comparative analysis. *Environmental Sciences Europe,* 35(1), 107. https://doi.org/10.1186/s12302-023-00817-1

Stanescu, S. G., Cucui, I., Ionescu, C. A., Paschia, L., Coman, M. D., Nicolau, N. L. G., ... , & Lixandru, M. L. (2021). Conceptual model for integrating environmental impact in managerial accounting information systems. *International Journal of Environmental Research and Public Health,* 18(4), 1791. https://doi.org/10.3390/ijerph18041791

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. Academy of Management Review, 20(3), 571–610. https://doi.org/10.5465/amr.1995.9508080331

Sulich, A. , & Kozar, L. J. (2024). Green competences: A review and future research in the context of green human resource management. *Economics and Environment*, 89(2), 713. https://doi.org/10.34659/eis.2024.89.2.713

Sun, H., Bahizire, G. M., Pea-Assounga, J. B. B. , & Chen, T. (2024). Enhancing employee green performance through green training: the mediating influence of organizational green culture and work ethic in the mining sector. *Journal of Cleaner Production*, 449, Article 141105. https://doi.org/10.1016/j.jclepro.2024.141105

United Nations Division for Sustainable Development. (UN DSD).(2001) . Environmental management accounting procedures and principles. United Nations. New York, United States. https://digitallibrary.un.org/record/446303

Van Doorn, J., Risselada, H. , & Verhoef, P. C. (2021). Does sustainability sell? The impact of sustainability claims on the success of national brands' new product introductions. *Journal of Business Research*, 137, 182-193. https://doi.org/10.1016/j.jbusres.2021.08.032

Von Der Gracht, H. A. (2012). Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technological Forecasting and Social Change*, 79(8), 1525-1536. https://doi.org/10.1016/j.techfore.2012.04.013

Voukkali, I., Papamichael, I., Loizia, P., Lekkas, D. F., Rodríguez-Espinosa, T., Navarro-Pedreño, J. , & Zorpas, A. A. (2023). Waste metrics in the framework of circular economy. *Waste Management & Research*, 41(12), 1741-1753. https://doi.org/10.1177/0734242X231190794

uncorrected proof

Wang, L. , & Si, Y. (2025). Analysis of the inverted “U” relationship between R&D intensity and green innovation performance: A study based on listed manufacturing enterprises in China. *Sustainability*, 17(17), 7625. https://doi.org/10.3390/su17177625

Weberová, D., Štarchoň, P., Juříková, M., Káčerková, E. , & Ližbetinová, L. (2024). Consumer preferences for eco-friendly products. *International Journal of Services, Economics and Management,* 15(4), 436-454. https://doi.org/10.1504/IJSEM.2024.139737

Wu, C. H. , & Fang, W. C. (2011). Combining the Fuzzy Analytic Hierarchy Process and the fuzzy Delphi method for developing critical competences of electronic commerce professional managers. *Quality & Quantity*, 45(4), 751-768. https://doi.org/10.1007/s11135-010-9425-6

Wu, Y. (2025). Financial and environmental impact in closed-loop circular supply chain. *Supply Chain Forum: An International Journal,* 26(4), 484–498. https://doi.org/10.1080/16258312.2025.2499507

Xing, K., Yang, F., Liu, P., Wang, J. , & Wu, J. (2025). Environmental violations and financial distress risk: Evidence from Chinese listed heavily polluting companies. *Pacific-Basin Finance Journal*, 89, Article 102583. https://doi.org/10.1016/j.pacfin.2024.102583

Xu, X., Pan, J. , & Meng, X. (2024). Intelligent manufacturing and green innovation—evidence from China’s listed manufacturing firms. *Sustainability*, 16(23), Article 10376. https://doi.org/10.3390/su162310376

Yafi, E., Tehseen, S. , & Haider, S. A. (2021). Impact of green training on environmental performance through mediating role of competencies and motivation. *Sustainability*, 13(10), 5624. https://doi.org/10.3390/su13105624

Yager, R. R. (1981). A new methodology for ordinal multiobjective decisions based on fuzzy sets. *Decision Sciences,* 12(4), 589–600. https://doi.org/10.1111/j.1540-5915.1981.tb00111.x

Yan, X., Bao, X., Zhao, R. , & Li, F. (2022). Performance measurement for green supplier selection based on data envelopment analysis. *Environmental Science and Pollution Research,* 29(30), 45960-45970. https://doi.org/10.1007/s11356-021-17897-2

Zheng, C., Xiao, F., Zeng, C. , & Yang, S. (2024). A pathway towards corporate sustainability: Does media attention matter?. *Heliyon*, 10(14), Article e34489. https://doi.org/10.1016/j.heliyon.2024.e34489

Uncorrected proof

Zhou, J., Sawyer, L. , & Safi, A. (2021). Institutional pressure and green product success: the role of green transformational leadership, green innovation, and green brand image. *Frontiers in Psychology*, 12, Article 704855. https://doi.org/10.3389/fpsyg.2021.704855

Zhou, P., Zhou, P., Yüksel, S., Dinçer, H. , & Uluer, G. S. (2020). Balanced scorecard-based evaluation of sustainable energy investment projects with IT2 fuzzy hybrid decision making approach. *Energies,* 13(1), 82. https://doi.org/10.3390/en13010082

uncorrected proof

Appendix A

In the Name of God

Esteemed Professor

Greetings and Respectfully,

You are hereby invited to review the attached questionnaire and provide your valuable opinions regarding the "Design of a Conceptual Model for the Green Balanced Scorecard (GBSC)," which has been selected as the topic of my master's thesis in accounting. The purpose of this research is to design a comprehensive and practical model for evaluating organizational performance from the perspective of sustainability and social responsibility. To this end, a questionnaire has been designed to assess the different dimensions of this model. Please review the questionnaire carefully and answer the questions presented. Your comments and suggestions will be highly effective in improving the quality of this research.

Thank you for your valuable cooperation.

Questionnaire 1 - Identifying Green Components in the Financial Perspective of the Balanced Scorecard

Question Number	Component Number	Identified Green Components	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1	1	Green product and service revenue (eco-products and services and the proportion of sales of environmentally friendly products)					
2	2	Green ancillary revenues (from recycling manufactured products and expanding revenue					

uncorrected proof

		opportunities from environmental activities)					
3	3	Environmental non-compliance costs (pollution-related fines and costs resulting from failure to address environmental activities)					
4	4	Environmental operating costs (disposal costs and costs of environmental initiatives)					
5	5	Resource-consumption optimization (maximizing energy savings and cost savings from environmental improvement)					
6	6	Waste-cost management (reducing operating costs by reducing waste and financial					

Uncorrected proof

		participation in reducing potential waste)					
7	7	Investment in sustainable technologies (technologies effective in conserving natural resources such as water and energy)					
8	8	Investment in the environmental sector					
9	9	Return on green investment (environmental efficiency and return on environmental investment)					
10	10	Increased profitability through a product-differentiation strategy with a green orientation					

uncorrected proof

Questionnaire 2 - Identifying Green Components in the Customer Perspective of the Balanced Scorecard

Question Number	Component Number	Identified Green Components	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1	11	Provision of environmentally friendly products (environmentally safe products, provision of clean products to customers, and the value of eco-products)					
2	12	Sustainable product design and packaging (green product design and packaging)					
3	13	Green brand and corporate environmental image (presenting the image of an environmentally responsible company and the brand reputation of eco-products)					

uncorrected proof

4	14	Enhanced media coverage from a green perspective					
5	15	Environmental risk management and legal compliance (reducing environmental risks and reducing lawsuits and warnings arising from environmental violations)					
6	16	Strength of regulatory relationships / compliance with future environmental regulations					
7	17	Green customer satisfaction (satisfaction and value creation for green customers)					

uncorrected proof

Questionnaire 3 - Identifying Green Components in the Internal-Process Perspective of the Balanced Scorecard

Question Number	Component Number	Identified Components	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1	18	Green manufacturing process (improving the environmental performance of production processes, adopting new and green technologies, reducing waste, and improving operational efficiency with a green approach)					
2	19	Sustainable energy management (consumption of renewable energy and energy efficiency)					
3	20	Resource efficiency					
4	21	Recycled-material management (use of recycled materials in production and the recycling rate for production and office materials)					
5	22	Implementation of a monitoring system for environmental pollutant emissions and the type of fuel consumed					
6	23	Reduction of production-generated wastewater					
7	24	Reduction of greenhouse gas emissions					
8	25	Waste management					

uncorrected proof

9	26	Reduction in the use of hazardous materials					
10	27	Use of a green supply chain (green supplier selection and relationships with sustainable suppliers)					
11	28	Use of environmental auditing and assessment to manage and control supplier performance					
12	29	Assessment of supplier environmental performance					
13	30	Environmentally friendly packaging					
14	31	Green product delivery					
15	32	Recycling and extension of the raw-material cycle					
16	33	Product reuse rate					

uncorrected proof

Questionnaire 4 - Identifying Green Components in the Learning and Growth Perspective of the Balanced Scorecard

Question Number	Component Number	Identified Green Components	Strongly Agree	Agree	No Opinion	Disagree	Strongly Disagree
1	34	Development of employees' environmental knowledge (increasing environmental awareness and employees' knowledge of environmentally friendly principles and environmental protection)					
2	35	Employees' environmental skills					

3	36	Number of hours of environmental training or consultation per employee/member					
4	37	Access to information and sharing of Green Knowledge					
5	38	Technological innovations and applied research related to product planning, packaging, and production processes to solve environmental problems					
6	39	Development of environmentally friendly materials					
7	40	Research and development for green technologies					
8	41	Hours allocated to research and development in environmental design					
9	42	Development of environmental infrastructure					
10	43	Environmental information systems					
11	44	Internal performance evaluation system incorporating environmental factors					
12	45	Number of environmental compliance audit programs for employees					
13	46	Encouraging employees toward environmental objectives					
14	47	Suggestions for environmental improvement					

Note: The wording of items 18, 34, 35 and 46 reflects the original Round-1 questionnaire. In the case of item 18, only its environmental component was retained in the analysis; the operational elements of safety and quality control were excluded at the extraction stage, as reported in

Uncorrected proof

Appendix C. The revised definitions of these indicators, as used in the analysis, are given in Appendix B.

Appendix B. The 47 indicators extracted through document analysis and their supporting sources

No.	code	Perspective	Component Name	Source
1	F1	Financial (F)	Green product and service revenue (eco-products and services and the proportion of sales of environmentally friendly products)	Klausmann et al. (2024) Cao et al. (2026)
2	F2	Financial (F)	Green ancillary revenues (from recycling manufactured products and expanding revenue opportunities from environmental activities)	Kosta and Ramadani (2025) Wu (2025)
3	F3	Financial (F)	Environmental non-compliance costs (pollution-related fines and costs resulting from failure to address environmental activities)	Xing et al. (2025) Shevchenko (2021)
4	F4	Financial (F)	Environmental operating costs (disposal costs and costs of environmental initiatives)	Nazir and Tennojiya (2026) Karim et al. (2025)
5	F5	Financial (F)	Resource-consumption optimization (maximizing energy savings and cost savings from environmental improvement)	Dhanorkar and Siemsen (2021) Panagiotopoulou et al. (2022)
6	F6	Financial (F)	Waste-cost management (reducing operating costs by reducing waste and financial participation in reducing potential waste)	Gull et al. (2022) Orazalin et al. (2026)
7	F7	Financial (F)	Investment in sustainable technologies	García-Quevedo and Jové-Llopis (2021)

uncorrected proof

			(technologies effective in conserving natural resources such as water and energy)	
8	F8	Financial (F)	Investment in the environmental sector	Cristea et al. (2024) Hidayat et al. (2024) Moussa and Elmarzouky (2023)
9	F9	Financial (F)	Return on green investment (environmental efficiency and return on environmental investment)	Chen and Ma (2021) Rahat and Nguyen (2022)
10	F10	Financial (F)	Increased profitability through a product-differentiation strategy with a green orientation	França et al. (2023) Alghababsheh et al. (2022)
11	C1	Customer (C)	Provision of environmentally friendly products (environmentally safe products, provision of clean products to customers, and the value of eco-products)	Morea et al. (2023) Weberová et al. (2024)
12	C2	Customer (C)	Sustainable product design and packaging (green product design and packaging)	Pogačar and Gregor-Svetec (2025) Li and Li (2024)
13	C3	Customer (C)	Green brand and corporate environmental image (presenting the image of an environmentally responsible company and the brand reputation of eco-products)	Ha (2022) Khan et al. (2022b)
14	C4	Customer (C)	Enhanced media coverage from a green perspective	Zheng et al. (2024) Chai et al. (2023) Pajuelo-Moreno et al. (2024)

uncorrected proof

15	C5	Customer (C)	Environmental risk management and legal compliance (reducing environmental risks and reducing lawsuits and warnings arising from environmental violations)	Du et al. (2024) Saeed and Cek (2025) Andreoli et al. (2025)
16	C6	Customer (C)	Strength of regulatory relationships / compliance with future environmental regulations	Sendawula et al. (2024) Chen et al. (2023)
17	C7	Customer (C)	Green customer satisfaction (satisfaction and value creation for green customers)	Martín-Gómez et al. (2024)
18	IP1	Internal Processes (IP)	Green manufacturing process (improving the environmental performance of production processes, adopting new and green technologies, reducing waste, and improving operational efficiency with a green approach)	Khan et al. (2022a) Bendig et al. (2023)
19	IP2	Internal Processes (IP)	Sustainable energy management (consumption of renewable energy and energy efficiency)	ISO 50001 (2018) Zhou et al. (2021) Sherif et al. (2022)
20	IP3	Internal Processes (IP)	Resource efficiency	Moraga et al. (2022) Aljamal et al. (2024)
21	IP4	Internal Processes (IP)	Recycled-material management (use of recycled materials in production and the recycling rate for production and office materials)	Dierkes and Siepelmeyer (2025) Aljamal et al. (2024)
22	IP5	Internal Processes (IP)	Implementation of a monitoring system for environmental pollutant emissions and the type of fuel consumed	ISO 14001(2015) Shi et al. (2023)

uncorrected proof

23	IP6	Internal Processes (IP)	Reduction of production-generated wastewater	GRI 303 (2021) Sherif et al. (2022)
24	IP7	Internal Processes (IP)	Reduction of greenhouse gas emissions	GRI 305 (2021) Panagiotopoulou et al. (2022)
25	IP8	Internal Processes (IP)	Waste management	GRI 306 (2020) Voukkali et al. (2023)
26	IP9	Internal Processes (IP)	Reduction in the use of hazardous materials	ISO 14001(2015) Slunge et al. (2023)
27	IP10	Internal Processes (IP)	Use of a green supply chain (green supplier selection and relationships with sustainable suppliers)	Abuzaid et al. (2024) Li and Zhong (2024)
28	IP11	Internal Processes (IP)	Use of environmental auditing and assessment to manage and control supplier performance	Amoako et al. (2023) Yan et al. (2022)
29	IP12	Internal Processes (IP)	Assessment of supplier environmental performance	Yan et al. (2022) Lee and Wu (2023)
30	IP13	Internal Processes (IP)	Environmentally friendly packaging	Postacchini et al. (2021) Weberová et al. (2024)
31	IP14	Internal Processes (IP)	Green product delivery	Sallnäs and Björklund (2023) Sarjana (2023)
32	IP15	Internal Processes (IP)	Recycling and extension of the raw-material cycle	Aljamal et al. (2024) Farooque et al. (2019)
33	IP16	Internal Processes (IP)	Product reuse rate	Betts et al. (2022) Mastria et al. (2024)
34	LG1	Learning and Growth (LG)	Development of employees' knowledge of the organization's environmental impacts and environmental management procedures	Dzhengiz and Niesten (2020) Amoako et al. (2023)
35	LG2	Learning and Growth (LG)	Employees' environment-specific technical skills	Sulich and Kozar (2024) Sun et al. (2024)
36	LG3	Learning and Growth (LG)	Number of hours of environmental training or consultation per employee/member	De Kruijf et al. (2025) Yafi et al. (2021)

Uncorrected proof

37	LG4	Learning and Growth (LG)	Access to information and sharing of green knowledge	Bresciani et al. (2023) Khan et al. (2024)
38	LG5	Learning and Growth (LG)	Technological innovations and applied research related to product planning, packaging, and production processes to solve environmental problems	D'Almeida and de Albuquerque (2024) Xu et al. (2024)
39	LG6	Learning and Growth (LG)	Development of environmentally friendly materials	Delaney et al. (2022) Martín-Gómez et al. (2024)
40	LG7	Learning and Growth (LG)	Research and development for green technologies	Ren and Li (2022) Saeedi and Ashraf (2024)
41	LG8	Learning and Growth (LG)	Hours allocated to research and development in environmental design	Camañes et al. (2024) Chang and Tsai (2015) Wang and Si (2025)
42	LG9	Learning and Growth (LG)	Development of environmental infrastructure	Jezzini et al. (2024) Khalili et al. (2024)
43	LG10	Learning and Growth (LG)	Environmental information systems	Mat Nawi et al. (2025) Stanescu et al. (2021)
44	LG11	Learning and Growth (LG)	Internal performance evaluation system incorporating environmental factors	Jassem et al. (2022) Sherif et al. (2022)
45	LG12	Learning and Growth (LG)	Number of environmental compliance audit programs for employees	Al Hanini (2021) Earnhart and Harrington (2021)
46	LG13	Learning and Growth (LG)	Linking employee incentives to environmental objectives	Chen et al. (2021) Kuo et al. (2022)
47	LG14	Learning and Growth (LG)	Suggestions for environmental improvement	Gusmerotti et al. (2023) Ma and Wang (2024)

Uncorrected proof

Appendix C. Decision rules applied in reducing the 76 extracted components to 47 indicators

Rule	Condition for removal or merger	Components	Examples
1	Verbatim duplication of a component already extracted from another source	7	"Share of sales of eco-friendly products" merged into F1
2	Conceptual overlap with an existing component; the two were merged under a single definition	19	"Energy efficiency" merged into IP2
3	Absence of an identifiable unit of measurement	0	-
4	Falling outside the environmental boundary of the GBSC as defined in Section 2.3	3	Effective maintenance; quality-control effectiveness; occupational safety
	Total removed or merged	29	

Note: The three components excluded under rule (iv) were judged to belong to general operational management or to the social pillar of sustainability rather than to environmental performance, consistent with the GBSC/SBSC distinction developed in Section 2. The content analysis was conducted by a single coder and reviewed by the supervising author; no inter-coder reliability coefficient was computed, and this is acknowledged in Section 6.

Appendix D. Comparison of Round-2 completers and non-respondents: non-response analysis

Panel A. Round-1 normalised scores by perspective				
Perspective	Completers Mean (n = 10)	Non-respondents Mean (n = 5)	Mann-Whitney U	Cliff's Delta
Financial (F)	0.868	0.870	23.5	-0.06
Customer (C)	0.825	0.857	24.0	+0.04
Internal processes (IP)	0.858	0.909	17.0	+0.32

uncorrected proof

Learning and growth (LG)	0.836	0.893	16.0	+0.36
Panel B. Panel characteristics (descriptive)				
Characteristic	Completers (n = 10)		Non-respondents (n = 5)	
Accounting	1		4	
Management	4		1	
Agricultural economics	5		0	
Professor or Associate Professor	3		4	
Assistant Professor or lecturer	7		1	
Teaching experience 20+ years	3		3	
Familiarity with GBSC: very good	7		5	
Panel C. Robustness of the Round-1 output				
Spearman's rho between the Round-1 defuzzified scores of the 47 indicators computed on the full panel (N = 15) and on completers only (n = 10)				0.932
Highest rank attained by any of the eight indicators removed after Round 1, when the scores are recomputed on completers only (1 = lowest score; 47 = highest)				17

Note: Panel A reports two-sided Mann-Whitney U tests (n1 = 5, n2 = 10; at the 0.05 significance level the critical value is U = 8, so any U greater than 8 indicates the absence of a significant difference). None of the four comparisons was significant; effect sizes were negligible for the financial and customer perspectives and small to moderate for the internal-process and learning and growth perspectives. Panel B is reported descriptively because the expected cell counts are too small for inferential testing; all three disciplines remained represented in Round 2, although their proportions changed. In Panel C the eight indicators removed after Round 1 all remain within the lowest third of the distribution when recomputed on the completers alone; because removal in Round 1 was governed by the three acceptance criteria jointly rather than by the defuzzified score alone, exact rank correspondence is not expected.

uncorrected proof