



Article

Developing a sustainable financial performance measurement instrument: Initial content-validity evidence for ESG practices, digital transformation, management innovation, and circular economy practices

Issa Al Harmali*, Noor Hidayah Abu, Azim Azuan Osman

College of Business, University of Utara Malaysia, Malaysia

*Corresponding author: Issa Al Harmali, alharmali99@hotmail.com

CITATION

Al Harmali I, Abu NH, Osman AA. (2026). Developing a sustainable financial performance measurement instrument: Initial content-validity evidence for ESG practices, digital transformation, management innovation, and circular economy practices. *Human Resources Management and Services*. 8(3): 6303. <https://doi.org/10.18282/hrms6303>

COPYRIGHT



Copyright © 2026 by author(s).

Human Resources Management and Services is published by Piscomed Publishing Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. <https://creativecommons.org/licenses/by/4.0/>

Abstract: This study develops a measurement instrument for assessing sustainable financial performance (SFP) through environmental, social, and governance (ESG) practices, digital transformation (DT), management innovation (MI), and circular economy practices (CEPs), and reports initial evidence of its content validity. The instrument is intended for Omani state-owned enterprises and is designed to capture organizational capabilities relevant to long-term financial sustainability, workforce readiness, leadership, innovation, responsible resource use, and organizational transformation. Six academic and professional experts evaluated item relevance, representativeness, and clarity using the Content Validity Index (CVI) and interrater agreement (IRA). Most items demonstrated strong expert agreement, while weaker items, particularly within MI and selected SFP-related dimensions, were identified for revision rather than treated as fully established measures. The findings therefore support progression to pilot testing but do not constitute evidence of construct validity, reliability, convergent validity, discriminant validity, or predictive validity. The study contributes an initial, context-specific measurement framework that can support subsequent psychometric validation and offers HR practitioners a structured basis for identifying capability, training, leadership, and change-management needs associated with ESG and digital transformation.

Keywords: ESG practices, digital transformation, sustainable financial performance, sustainable social development, Oman Vision 2040, state-owned enterprises

1. Introduction

Sustainable development has become a central priority for governments, businesses, and international organizations as they strive to achieve economic prosperity while safeguarding environmental resources, promoting social well-being, and strengthening institutional governance. This global agenda has been reinforced by the United Nations Sustainable Development Goals (SDGs), which encourage organizations to integrate sustainability into their strategic and operational activities to create long-term value for both stakeholders and society. As a result, organizations are increasingly expected to generate not only financial returns but also positive social and environmental outcomes that contribute to inclusive and sustainable development.

Within this context, sustainable financial performance (SFP) has become an increasingly important topic as organizations seek to balance financial success with broader sustainability objectives (Moussa et al., 2024). It reflects an organization's ability to achieve consistent financial performance while integrating environmental, social, and governance (ESG) considerations into its strategic planning and day-to-day operations (Chung et al., 2021). Unlike traditional financial performance, which often emphasizes short-term profitability, SFP focuses on creating long-term value through financial resilience, responsible

resource management, and sustainable business practices (Gong et al., 2019). When sustainability is embedded within organizational strategy rather than treated as a separate initiative, it strengthens long-term competitiveness and supports wider national development goals (Fernando et al., 2019; Gupta et al., 2021). In this way, sustainable financial performance not only benefits organizations but also contributes to sustainable social development by promoting responsible governance, resource efficiency, and long-term economic resilience.

Evidence from leading international utility companies demonstrates that sustainability and strong financial performance can be achieved simultaneously. For instance, Veolia generated €45.35 billion in revenue in 2023 while committing to a 50% reduction in greenhouse gas emissions by 2032. Similarly, Suez enhanced its environmental performance alongside a €100 million investment in water infrastructure projects in Angola (Europa Wire, 2024; Waste Dive, 2024; Suez, 2024). These examples illustrate how sustainability initiatives can strengthen organizational performance while delivering broader environmental and social benefits. In Oman, these principles are reflected in Oman Vision 2040, which aims to build a diversified, knowledge-based economy driven by innovation, sustainability, and private-sector participation (Ministry of Economy, 2025). The vision also promotes the integration of ESG principles across both public and private institutions, with strategic priorities including renewable energy, efficient resource management, innovation, and economic diversification (Times News Service, 2022; Oman Vision 2040, 2024; World Bank, 2025). These priorities are closely aligned with the Sustainable Development Goals (SDGs), particularly SDG 8 (Decent Work and Economic Growth), SDG 9 (Industry, Innovation and Infrastructure), SDG 12 (Responsible Consumption and Production), and SDG 16 (Peace, Justice and Strong Institutions).

The Oman Investment Authority (OIA), established in 2020, plays a strategic role in advancing these national priorities by managing a diversified portfolio of state-owned enterprises and attracting investments that support economic diversification (Al Murshidi, 2025). Beyond generating financial returns, OIA companies are expected to contribute to public value through sound governance, transparency, and sustainable business practices. Nevertheless, several governance-related challenges continue to affect public-sector organizations, including limited ownership transparency and weaknesses in accountability mechanisms (Devdiscourse, 2024). The International Monetary Fund (2024) further reported that transparency remains one of the most significant concerns facing OIA companies, with more than 80% of analysts identifying governance and disclosure as major areas requiring improvement. These challenges highlight the importance of developing reliable measurement instruments that can evaluate organizational capabilities supporting both sustainable financial performance and broader sustainable social development.

Against this background, this study develops a self-administered questionnaire (SAQ) and evaluates its initial content validity through expert review. The research gap is not simply the absence of measures for ESG, DT, MI, CEPs, or SFP individually; rather, existing measures are commonly applied as separate scales and provide limited context-specific guidance for assessing how these organizational capabilities can be examined together in Omani state-owned enterprises. The proposed instrument therefore brings the constructs into one coordinated measurement framework while retaining them as theoretically distinct constructs. For HRM scholarship, this framework is relevant because implementation of ESG and digital transformation depends on workforce competencies, sustainability-oriented training, leadership development, innovation-supportive culture, talent planning, and change readiness. The present

article is limited to the instrument-development and content-validation stage. It does not test causal relationships among the constructs. Its immediate contribution is to establish the conceptual content and expert-assessed clarity of the items before pilot testing and subsequent empirical psychometric validation.

2. Theoretical and conceptual framework

The instrument is organized around five theoretically distinct constructs: ESG practices, digital transformation (DT), management innovation (MI), circular economy practices (CEPs), and sustainable financial performance (SFP). In the broader research model, ESG practices and DT are treated as antecedent organizational capabilities, MI and CEPs represent organizational mechanisms through which sustainability- and technology-oriented strategies may be translated into organizational outcomes, and SFP is the focal outcome construct. These proposed relationships provide the conceptual logic for the instrument; however, they are not tested in the present content-validation study.

The framework is consistent with a stakeholder perspective, which emphasizes organizational responsibilities to multiple stakeholder groups; a resource-based perspective, which views organizational capabilities and knowledge as sources of sustained performance; and a dynamic-capabilities perspective, which emphasizes the ability to reconfigure resources and routines in response to technological and sustainability-related change. Within HRM, these perspectives also draw attention to human capabilities: ESG implementation requires sustainability competencies and responsible leadership; DT requires digital skills and workforce adaptability; MI depends on leadership, learning, and an innovation-supportive culture; and CEPs require employees to adopt resource-efficient and circular work practices.

Conceptually, the subsequent empirical phase will examine ESG and DT as antecedents of SFP and will assess the mediating roles of MI and CEPs. The present study does not infer causality from expert ratings. Instead, expert review is used only to determine whether the proposed items adequately represent and clearly express the intended construct domains. This distinction separates the current content-validity stage from later testing of the measurement and structural models.

3. Materials and methods

3.1 Instrument development

The questionnaire was developed through a construct-by-construct process. The domains of SFP, ESG practices, DT, MI, and CEPs were first defined from the theoretical and empirical literature used in the study. Candidate items were then adapted and consolidated to reflect the organizational context of OIA-affiliated state-owned enterprises. Items were screened for conceptual relevance, duplication, breadth, clarity, and suitability for managerial respondents. The purpose of this stage was to achieve adequate domain coverage rather than to claim that the resulting items already possessed psychometric reliability or construct validity. Expert review subsequently served as an independent content-validation stage in which the wording and representativeness of the candidate items were evaluated before pilot testing.

3.2 Expert selection and review procedure

Six experts were purposively selected to provide complementary academic and practitioner perspectives. Eligibility required relevant academic expertise in management or measurement and/or professional experience in sustainability, ESG, investment, circular economy, or organizational management. The panel included senior sustainability and ESG practitioners as well as university academics. Individuals without relevant subject-matter or methodological

experience were not included. Diversity of role, sector experience, education, and professional background was used to reduce dependence on a single perspective. Experts completed their assessments independently, and their quantitative ratings were considered together with written qualitative comments. Potential bias was further limited by reporting the expert profile and by documenting substantive comments and the actions taken in response.

Content validation is a foundational step in survey instrument development that ensures that the items included in a questionnaire accurately represent the theoretical constructs being measured (Lambert & Newman, 2023). Content validity focuses on the degree to which individual items and the overall instrument representatively sample the full domain of the construct of interest (Almanasreh et al., 2019). In other words, ensuring content validity means confirming that all significant facets of a construct, such as ESG practices, digital transformation, or management innovation, are adequately covered by the survey items. Researchers commonly use expert judgment, systematic item reviews, and quantitative indices such as the content validity index (CVI) to establish this form of validity (Polit & Beck, 2006). Mokkink et al. (2025) explained that in the absence of rigorous content validation, a measurement instrument may yield misleading or invalid results, even if other psychometric properties are satisfactory. Thus, content validation strengthens the appropriateness, relevance, and comprehensiveness of questionnaire items before pilot testing and full-scale data collection (Kim, 2009).

Although content validity has traditionally been viewed as relying heavily on expert judgment and therefore subject to potential subjectivity, recent methodological research indicates that this limitation can be addressed through structured and transparent quantitative procedures (Almanasreh et al., 2019; Lim, 2024). In particular, the content validity index (CVI) enables expert assessments to be converted into numerical indicators of item relevance, clarity, and representativeness, thereby strengthening the rigor of the validation process (Rubio et al., 2003; Mokkink et al., 2025). Accordingly, the present study applied CVI procedures to validate the measurement items developed for Environmental, Social, and Governance (ESG) practices, Digital Transformation (DT), Management Innovation (MI), Circular Economy Practices (CEPs), and Sustainable Financial Performance (SFP) within Oman Investment Authority-affiliated organizations. Given the multidimensional and strategically interconnected nature of these constructs, CVI-based expert evaluation is particularly important to ensure that each item adequately represents its intended domain while minimizing construct overlap and ambiguity (Lim, 2024). Furthermore, documenting content validity evidence through CVI is consistent with internationally recognized measurement standards and best practices endorsed by leading professional bodies, including the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education (Plake & Wise, 2014).

In addition to assessing item relevance through the content validity index (CVI), methodological scholars recommend examining interrater agreement (IRA) to evaluate the consistency and reliability of expert judgments during the content validation process (Rubio et al., 2003). IRA analysis determines the degree to which experts provide consistent ratings across items within each construct, thereby strengthening confidence in the objectivity of the validation outcomes. Technically, IRA is calculated by dividing the total number of agreed ratings across items by the total number of items within a given dimension, allowing researchers to assess agreement at the construct level rather than relying solely on individual item evaluations. In the context of the present study, which involves multidimensional constructs such as ESG practices, digital transformation (DT), management innovation (MI), circular economy practices (CEPs), and sustainable

financial performance (SFP), assessing IRA was particularly important to ensure that expert evaluations were stable and consistent across diverse but conceptually related domains. Accordingly, this study adhered to the CVI and IRA reporting procedures proposed by Rubio et al. (2003) to provide robust and transparent evidence of content validity for the questionnaire items, thereby reinforcing the methodological rigor of the instrument prior to pilot testing and empirical analysis.

For this study, an expert rating of 3 or 4 on the four-point scale was treated as endorsement of an item for the relevant criterion. I-CVI was calculated as the number of experts providing an endorsed rating divided by the total number of experts ($n = 6$). Construct-level CVI was summarized across the items within each construct. IRA was used as a complementary indicator of consistency in expert judgments and was reported at item and construct levels. A value of 0.80 was used as the reference criterion for adequate agreement. Accordingly, values below 0.80 are described in the revised manuscript as indicating weaker or insufficient agreement requiring item review, rather than as automatically acceptable.

3.3 Planned empirical validation

Following revision of the content-validated item pool, the next phase will involve pilot testing and empirical validation with managers from OIA-affiliated entities. The planned minimum sample is 129 respondents, based on an a priori power-analysis target for the broader model. Respondents will be drawn from middle and senior management positions because the constructs require organizational-level knowledge of ESG, digital transformation, innovation, circular-economy practices, and financial performance. The empirical phase will examine internal consistency and measurement-model properties, including indicator reliability, convergent validity, and discriminant validity, before testing the proposed structural relationships and mediation effects. These analyses are outside the scope of the present expert-based content-validation study.

4. Results

Validity is widely recognized as a foundational requirement for ensuring that research instruments accurately and meaningfully capture the constructs they are intended to measure. Lim (2024) conceptualizes validity not as a single statistical outcome but as a cumulative and multistage process that evolves throughout the research lifecycle. This perspective emphasizes that validity evidence should be built progressively, starting from conceptual clarity during instrument design and extending to theory testing and outcome prediction. At the initial stage, content validity and face validity play critical roles (Polit & Beck, 2006).

Table 1. Profile of the Experts and Practitioners.

Expert Code	Gender	Education	Designation (Organization)	Relevance
Expert #1	Male	PhD	VP-Sustainability and Circular Economy	More than 20 years' experience in Sustainability and Circular Economy
Expert #2	Male	Masters	Sustainability Manager	15 years' experience in Sustainability
Expert #3	Male	Masters	Head –Investments & ESG	9 years' experience in Sustainability and ESG
Expert #4	Female	PhD	Senior Lecturer (UUM)	21 years' experience in STML teaching
Expert #5	Female	PhD	Senior Lecturer (UUM)	21 years' experience in STML teaching
Expert #6	Male	PhD	Assistant Professor (UTAS)	5 years' experience in UTAS teaching

Methodological research shows that including multiple experts in content validation strengthens the credibility of the evaluation (Yamanaka, 2025). Although earlier guidelines suggest that at least three experts are sufficient to

begin content validation, recent studies recommend involving a larger panel, typically between five and six experts, to improve the consistency and quality of judgments (Boateng et al., 2018). In addition to the number of experts, scholars emphasize the importance of having a diverse panel that includes individuals with different perspectives, such as subject-matter specialists, methodological experts, and professionals from the intended population who use or are affected by the instrument (Carpenter, 2018; Yang et al., 2025).

Following these recommendations, this study, as shown in **Table 1**, assembled a panel of six experts representing a range of expertise relevant to the research context. The panel included academic researchers with content and methodological experience, an industry consultant, and practitioners drawn from the target population. This composition satisfies the current standards for both the number and diversity of experts recommended in the literature. On the basis of the experts' comments and suggestions, the required corrections for the items were determined as presented in **Table 2**.

Table 2. Summary of suggestions from academics and experts.

Representative	Comments and suggestions
Expert #1	The instrument is theoretically sound and contextually relevant to Oman Vision 2040 but requires correction of a duplicated MI item and clearer instructions for the 7-point scale to enhance measurement quality.
Expert #2	The constructs are relevant and aligned with OIA and national ESG expectations, but several items are overly broad or compliance-focused, risking generic agreement and limiting differentiation, highlighting the need for clearer timeframes and stronger maturity- and performance-oriented measures.
Expert #3	The items are well designed, but the demographic section should be refined by removing the narrow employee range (1–10) and expanding the list of sustainability representatives to include administrative roles.
Expert #4	Provide clear explanations of key concepts (e.g., AI and digital services) and revise Question 7 to ensure it is complete and easily understood by respondents.
Expert #5	Overall, the instrument shows high clarity and representativeness, with only minor revisions needed, including small refinements to SP3 and EP3 and the removal or revision of the redundant MI2 item.
Expert #6	The survey questions are clear and easy to understand.

Following data collection, the focus shifts toward construct validity, particularly convergent and discriminant validity (Lambert & Newman, 2023). Convergent validity assesses whether indicators intended to measure the same construct are sufficiently correlated, whereas discriminant validity ensures that theoretically distinct constructs remain empirically separate (Lim, 2024). Establishing both forms of validity is essential to avoid construct overlaps and to ensure meaningful interpretation of structural relationships in multivariate analyses (Lambert & Newman, 2023).

Beyond measurement adequacy, higher-order validity evidence is needed to demonstrate theoretical and practical relevance (Lim, 2024). Likewise, nomological validity examines whether the relationships among constructs are consistent with established theoretical frameworks, thereby confirming that the measurement behaves as expected within a broader network of constructs. Predictive validity further extends this logic by assessing the ability of constructs to explain or predict relevant outcomes, reinforcing the applied value of the instrument (Lim, 2024).

Taken together, this layered view of validity underscores that rigorous empirical research requires more than post hoc statistical testing. Instead, an integrated validation strategy that combines expert evaluation, statistical evidence, and theoretical coherence is needed (Lim, 2024). Such an approach enhances the

credibility, interpretability, and applicability of research findings, particularly in studies relying on self-report instruments and latent constructs (Lim, 2024).

Table 3. Summary of expert representativeness and clarity ratings.

Item Code	Representativeness								Clarity							
	E #1	E #2	E #3	E #4	E #5	E #6	nr	na	E #1	E #2	E #3	E #4	E #5	E #6	nr	na
SFP01	4	3	4	4	4	3	6	4	4	4	4	4	4	3	6	5
SFP02	4	4	4	4	4	3	6	5	4	4	4	4	4	2	5	5
SFP03	4	4	4	4	4	2	5	5	4	4	4	4	4	3	6	5
SFP04	4	3	4	4	4	3	6	4	4	4	4	4	4	3	6	5
SFP05	4	3	4	4	4	2	5	4	3	4	4	4	4	2	5	4
SFP06	4	4	4	4	4	3	6	5	4	4	4	4	4	2	5	5
EP1	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
EP2	4	4	4	4	4	3	6	5	4	4	4	3	4	3	6	5
EP3	4	4	4	3	4	2	5	5	4	4	4	4	3	3	6	4
EP4	4	4	4	4	4	2	5	5	4	4	4	3	4	2	5	4
EP5	4	2	4	3	4	3	5	3	4	4	4	4	4	3	6	5
EP6	4	4	4	4	4	4	6	6	4	4	4	4	4	3	6	5
SP1	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
SP2	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
SP3	4	2	4	4	4	4	5	5	4	4	4	4	3	3	6	5
SP4	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
SP5	4	4	4	4	4	2	5	5	4	4	4	4	4	3	6	5
SP6	4	4	4	4	4	2	5	5	4	4	4	4	4	3	6	5
GP1	4	4	4	4	4	4	6	6	4	4	4	4	4	3	6	5
GP2	4	4	4	4	4	4	6	6	4	4	4	4	4	3	6	5
GP3	4	4	4	4	4	3	6	5	4	4	4	4	4	3	6	5
GP4	4	4	4	4	4	3	6	5	4	4	4	4	4	2	5	5
GP5	4	4	4	4	4	3	6	5	4	4	4	4	4	3	6	5
GP6	4	2	4	4	4	4	5	5	4	4	4	4	4	3	6	5
DT1	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
DT2	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
DT3	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
DT4	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
DT5	4	3	4	4	4	4	6	5	4	4	4	4	4	4	6	6
DT6	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
MI1	4	3	4	3	4	4	6	5	1	3	4	3	4	4	5	3
MI2	1	1	4	3	4	4	4	3	4	2	4	3	4	4	5	4
MI3	4	3	4	3	1	4	5	3	4	3	4	3	4	4	6	4
MI4	4	3	4	4	4	4	6	5	4	3	4	4	4	4	6	5
MI5	4	3	4	4	4	4	6	5	4	4	4	4	4	4	6	6
MI6	4	3	4	4	4	4	6	5	4	3	4	4	4	4	6	5
CEP1	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
CEP2	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
CEP3	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
CEP4	4	4	4	4	4	4	6	6	4	4	4	4	4	4	6	6
CEP5	4	2	4	4	4	4	5	5	4	4	4	4	4	4	6	6

Note. Nr = number of experts who rated 3 or 4, na = individual item's agreement

To ensure the adequacy and representativeness of the measurement instrument, content validity was assessed via the content validity index (CVI) in conjunction with interrater agreement (IRA) (Polit & Beck, 2006). A panel of subject-matter experts evaluated the relevance of each questionnaire item via a four-point ordinal scale ranging from “not representative” to “representative,” as recommended in the measurement validation literature (Polit et al., 2007) (**Table 3**). Indeed, consistent with established guidelines, ratings indicating minor revision or full representativeness were considered acceptable indicators of item relevance, and the item-level CVI was calculated as the proportion of experts endorsing each item relative to the total number of experts (Almanasreh et al., 2019).

In addition, interrater agreement was computed to assess the degree of consistency among expert judgments across items within each construct, thereby complementing the CVI results by capturing consensus among evaluators (Mokkink et al., 2025). As shown in **Table 3**, the combined application of CVI and IRA strengthens the methodological rigor of the instrument validation process by confirming both content relevance and expert agreement prior to pilot testing and full-scale data collection.

In line with established methodological recommendations, both the content validity index (CVI) and the interrater agreement (IRA) were examined at the item level and the construct (dimension) level to ensure a comprehensive assessment of the measurement instrument (Rubio et al., 2003). To clearly differentiate between these two levels of analysis and enhance clarity in reporting, the terminology proposed by Polit and Beck (2006) was adopted, whereby the I-CVI refers to the content validity of individual items and the S-CVI represents the overall content validity of each construct. Following the same logic, this study applied the terms I-IRA and S-IRA to denote interrater agreement at the item and construct levels, respectively.

This approach is particularly relevant given the multidimensional nature of the study variables: environmental, social, and governance (ESG) practices; digital transformation (DT); management innovation (MI); circular economy practices (CEPs); and sustainable financial performance (SFP), where both individual indicators and overall construct representations must be carefully evaluated. Consistent with prior measurement research, a CVI threshold of 0.80 was considered sufficient to establish acceptable content validity for survey-based instruments, whereas an IRA value of 0.80 or above was interpreted as evidence of adequate agreement and reliability among expert ratings (Davis, 1992; Lynn, 1986). All the item-level and construct-level CVI and IRA results are summarized in **Table 4**, which provides transparent and systematic evidence of the instrument's content validity and expert agreement prior to pilot testing and empirical analysis.

As shown in **Table 4**, for item representativeness, the majority of the items achieved I-CVI values of 1.00, indicating unanimous agreement among the experts. Several items had I-CVI values of 0.83, reflecting agreement by five out of six experts, whereas only a very limited number of items, mainly within the Management Innovation construct, had lower values of 0.67, suggesting the need for minor wording refinement. At the construct level, the S-CVI values for representativeness ranged from 0.91 to 1.00, demonstrating particularly strong content coverage.

With respect to item clarity, similar patterns were observed. Most of the items had I-CVI values between 0.83 and 1.00, indicating that the experts generally agreed that the items were clearly worded and easy to understand. A small number of items recorded clarity scores of 0.67, again primarily within the Management Innovation and Sustainable Financial Performance dimensions. At the construct level, the S-CVI values for clarity ranged from 0.92-1.00, confirming

that the overall wording and structure of the questionnaire were coherent across all dimensions.

The IRA results show substantial variation across constructs. Digital Transformation and Circular Economy Practices achieved the strongest construct-level agreement. Management Innovation produced the weakest S-IRA values (approximately 0.72–0.75), which fall below the 0.80 reference criterion used in this study. These values are therefore interpreted as evidence that the MI item set required substantive review rather than as acceptable agreement. The lower scores, together with the experts' qualitative comments, were used to identify potentially redundant, broad, or unclear items for revision before pilot testing. The combined CVI and IRA evidence supports progression of the revised item pool to the next development stage, but it does not establish reliability or construct validity.

Table 4. CVI and IRA results.

Dimensions	Item Code	Representativeness					Clarity					I-IRA	S-IRA
		Nr	I-CVI	S-CVI	na	I-IRA	S-IRA	nr	I-CVI	S-CVI	na		
Sustainable Financial Performance	SFP01	6	1		4	0.66		6	1		5	0.833	
	SFP02	6	1		5	0.833		5	0.833		5	0.833	
	SFP03	5	0.833	0.94	5	0.833	0.75	6	1	0.92	5	0.833	0.81
	SFP04	6	1		4	0.66		6	1		5	0.833	
	SFP05	5	0.833		4	0.66		5	0.833		4	0.66	
	SFP06	6	1		5	0.833		5	0.833		5	0.833	
Environmental Practices	EP1	6	1		6	1		6	1		6	1	
	EP2	6	1		5	0.833		6	1		5	0.833	
	EP3	5	0.833	0.92	5	0.833	0.833	6	1	0.972	4	0.66	0.81
	EP4	5	0.833		5	0.833		5	0.833		4	0.66	
	EP5	5	0.833		3	0.5		6	1		5	0.833	
	EP6	6	1		6	1		6	1		5	0.833	
Social Practices	SP1	6	1		6	1		6	1		6	1	
	SP2	6	1		6	1		6	1		6	1	
	SP3	5	0.833	0.92	5	0.833	0.916667	6	1	1	5	0.833	0.92
	SP4	6	1		6	1		6	1		6	1	
	SP5	5	0.833		5	0.833		6	1		5	0.833	
	SP6	5	0.833		5	0.833		6	1		5	0.833	
Governance Practices	GP1	6	1		6	1		6	1		5	0.833	
	GP2	6	1		6	1		6	1		5	0.833	
	GP3	6	1	0.97	5	0.833	0.88	6	1	0.97	5	0.833	0.833
	GP4	6	1		5	0.833		5	0.833		5	0.833	
	GP5	6	1		5	0.833		6	1		5	0.833	
	GP6	5	0.833		5	0.833		6	1		5	0.833	
Digital Transformation	DT1	6	1		6	1		6	1		6	1	
	DT2	6	1		6	1		6	1		6	1	
	DT3	6	1	1	6	1	0.97	6	1	1	6	1	1
	DT4	6	1		6	1		6	1		6	1	
	DT5	6	1		5	0.833		6	1		6	1	
	DT6	6	1		6	1		6	1		6	1	
Management Innovation	MI1	6	1		5	0.833		5	0.833		3	0.5	
	MI2	4	0.66		3	0.5		5	0.833		4	0.66	
	MI3	5	0.833	0.92	3	0.5	0.72	6	1	0.94	4	0.66	0.75
	MI4	6	1		5	0.833		6	1		5	0.833	

Continued table:

Dimensions	Item Code	Representativeness						Clarity					
		Nr	I-CVI	S-CVI	na	I-IRA	S-IRA	nr	I-CVI	S-CVI	na	I-IRA	S-IRA
Circular Economy Practices	MI5	6	1		5	0.833		6	1		6	1	
	MI6	6	1		5	0.833		6	1		5	0.833	
	CEP1	6	1		6	1		6	1		6	1	
	CEP2	6	1		6	1		6	1		6	1	
	CEP3	6	1	0.96	6	1	0.966	6	1	1	6	1	1
	CEP4	6	1		6	1		6	1		6	1	
	CEP5	5	0.833		5	0.833		6	1		6	1	

5. Discussion

This study examined the initial content validity of a measurement instrument designed to assess SFP alongside ESG practices, DT, MI, and CEPs. The findings indicate generally strong expert judgments of item relevance and clarity, but they should be interpreted as evidence from one early stage of instrument development. They do not demonstrate empirical reliability, construct validity, convergent validity, discriminant validity, nomological validity, or predictive validity. The principal value of the current results is therefore to identify which items are sufficiently clear and representative to progress to pilot testing and which require revision.

The high content validity scores indicate that the questionnaire items adequately represent their intended constructs, while the clarity ratings suggest that the items are generally easy to understand and interpret. This supports previous studies arguing that content validity is strengthened when measurement items accurately reflect theoretical concepts and are refined through systematic expert feedback (Rubio et al., 2003; Almanasreh et al., 2019). Although several experts recommended minor revisions to improve wording and remove overlapping items, these changes enhanced the overall quality of the instrument without altering its conceptual foundation.

Digital Transformation and Circular Economy Practices showed the strongest expert agreement, whereas Management Innovation showed weaker agreement. The MI S-IRA values below 0.80, together with item-level I-CVI values as low as 0.67, indicate that this construct required more than minor wording refinement. Expert feedback identified a potentially redundant MI item and highlighted the need to improve specificity and clarity. Selected SFP and ESG-related items also received comments concerning breadth, compliance orientation, or wording. These findings were treated as substantive item-development evidence: redundant content was removed or revised, unclear wording was refined, and items were retained for the next stage only where their conceptual contribution remained justified.

5.1 HRM implications

The framework has direct relevance to HRM because sustainable organizational performance depends on the capabilities and behaviors through which ESG, digital, innovation, and circular-economy strategies are implemented. HR professionals can use the revised item framework as an initial diagnostic structure to identify areas requiring sustainability-oriented training, digital-skills development, leadership development, workforce planning, and change-management support. For example, weak organizational responses to DT-related items may indicate capability gaps requiring reskilling, while MI-related weaknesses may point to leadership, learning, job-design, or cultural barriers to innovation. Similarly, ESG and CEP-related assessments can inform

competency frameworks, employee awareness programs, and HR practices that reinforce responsible and resource-efficient behavior. Because the instrument has not yet undergone empirical psychometric validation, these applications should presently be viewed as potential uses to be evaluated in subsequent research.

5.2 Theoretical and practical contribution

The study's contribution lies in developing a coordinated, context-specific item framework rather than claiming a new fully validated scale. Its conceptual novelty is the joint organization of five distinct constructs within an Omani state-owned-enterprise setting and the explicit connection of these constructs to workforce capabilities and organizational transformation. The content-validation evidence provides a defensible basis for pilot testing and for later examination of the proposed antecedent and mediating relationships. Subsequent research must determine whether the proposed structure is supported empirically and whether the instrument demonstrates adequate reliability and validity.

6. Conclusion and limitations

This study reports an initial content-validation stage in the development of a self-administered questionnaire covering ESG practices, DT, MI, CEPs, and SFP. Six experts with relevant academic and professional backgrounds evaluated the representativeness and clarity of the proposed items. CVI and IRA results, considered together with qualitative feedback, indicated that many items were suitable for progression while several weaker items, particularly within MI and selected SFP-related dimensions, required substantive revision.

The revised questionnaire should therefore be regarded as a content-reviewed instrument prepared for pilot testing, not as a fully validated or reliable measure. Its potential practical value lies in providing organizations and HR professionals with a structured way to examine sustainability-related capabilities, including leadership, employee competencies, digital readiness, innovation support, and resource-efficient practices. These potential applications require confirmation through empirical testing.

The study has several limitations. First, the evidence is based on expert judgment from a six-member panel and does not include data from the target respondent population. Second, reliability and construct validity have not yet been established. Third, lower agreement for several MI and SFP-related items shows that the item pool remains subject to refinement. Finally, the instrument was developed for the Omani state-owned-enterprise context, so its applicability to other sectors and countries cannot be assumed. Future research should conduct pilot testing and full empirical psychometric validation, assess internal consistency, convergent and discriminant validity, and test the proposed structural and mediation relationships using an adequately powered managerial sample.

Accordingly, the present contribution is methodological and preliminary: it establishes documented content-validity evidence and a clearer conceptual basis for the next phase of scale development. The revised instrument offers a foundation for future research on how ESG practices and digital transformation may relate to SFP through management innovation and circular economy practices, while also enabling future examination of the HR capabilities required to support these organizational transformations.

Author contributions: Issa Al Harmali conceived the study, designed the research methodology, developed the measurement instrument, conducted the expert validation process, analyzed the data, prepared the figures and tables, and wrote the first draft of the manuscript. Noor Hidayah Abu contributed to the study design, reviewed the methodology, critically revised the manuscript, and supervised the research. Azim Azuan Osman contributed to the conceptual development of the study, provided methodological guidance, reviewed and edited

the manuscript, and supervised the overall research project. All authors reviewed the manuscript, approved the final version, and agreed to be accountable for all aspects of the work.

Funding: This research received no external funding.

Ethical approval: This study was reviewed and approved by the Universiti Utara Malaysia Research Ethics Committee.

Informed consent statement: Informed consent was obtained from all experts who voluntarily participated in the content validation process before completing the evaluation of the questionnaire.

Data availability statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available because they include expert evaluation records collected for research purposes and are subject to confidentiality considerations.

Acknowledgments: The authors sincerely thank the academic experts and industry practitioners who generously contributed their time and expertise during the content validation process. Their constructive comments and valuable recommendations greatly improved the quality and clarity of the measurement instrument.

Conflict of interest: The authors declare that they have no known financial or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 15(2), 214–221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- [2] Al-Murshidi, S. (2025). Oman Investment Authority and economic diversification strategies in Oman Vision 2040. *Middle East Economic Review*, 17(2), 45–59.
- [3] Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>
- [4] Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, 12(1), 25–44. <https://doi.org/10.1080/19312458.2017.1396583>
- [5] Chen, A., Li, L., & Shahid, W. (2024). Digital transformation as the driving force for sustainable business performance: A moderated mediation model of market-driven business model innovation and digital leadership capabilities. *Heliyon*, 10(8), e28326. <https://doi.org/10.1016/j.heliyon.2024.e28326>
- [6] Chung, C. Y., Jung, S., & Young, J. (2021). Do ESG activities improve corporate financial performance? Evidence from global firms. *Sustainability*, 13(7), 3746. <https://doi.org/10.3390/su13073746>
- [7] Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194–197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- [8] Devdiscourse. (2024). Corporate governance challenges in Oman state-owned enterprises. <https://www.devdiscourse.com>
- [9] Fernando, Y., Jabbour, C. J. C., & Wah, W. X. (2019). Pursuing green growth in technology firms through sustainability strategy. *Business Strategy and the Environment*, 28(5), 902–915. <https://doi.org/10.1002/bse.2292>
- [10] Gong, B., Greenwood, R., & Song, C. X. (2019). Sustainable long-term value creation: A new finance focus for boards of directors. *Journal of Corporate Accounting & Finance*, 30(3), 25–35. <https://doi.org/10.1002/jcaf.22382>
- [11] Gupta, J., Crilly, D., & Greckhamer, T. (2021). Stakeholder engagement strategies, national context, and firm performance. *Business Strategy and the Environment*, 30(4), 1899–1916. <https://doi.org/10.1002/bse.2729>
- [12] Huang, C. H., & Huang, Y. C. (2024). Exploring the link between ESG performance and corporate sustainability outcomes. *Sustainability*, 16(3), 1420. <https://doi.org/10.3390/su16031420>
- [13] International Monetary Fund. (2024). Oman: Selected issues report. <https://www.imf.org>
- [14] Lambert, V. A., & Newman, E. (2023). Content validity in instrument development. *Journal of Nursing Measurement*, 31(1), 3–10.
- [15] Lavrakas, P. J. (2008). *Encyclopedia of survey research methods*. Sage Publications.
- [16] Lim, W. M. (2024). Toward a unified view of research validity. *Journal of Marketing Analytics*, 12(1), 1–12.
- [17] Lu, J., Song, Z., Gheith, M. H., Troise, C., & Bresciani, S. (2025). Internal dynamics of SMEs for their sustainable development: Investigating the antecedents and outcomes of sustainable innovation behavior and ESG strategies. *Business Strategy and the Environment*. Advance online publication. <https://doi.org/10.1002/bse.3845>
- [18] Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385.
- [19] Ministry of Economy. (2025). Oman Vision 2040 progress report. Government of Oman.
- [20] Mokkink, L. B., Prinsen, C. A. C., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Terwee, C. B. (2025). COSMIN methodology for content validity. *Quality of Life Research*, 34, 1–14.

- [21] Moussa, T., Allam, A., & El-Banna, S. (2024). Sustainable financial performance and ESG integration. *Sustainability Accounting, Management and Policy Journal*, 15(2), 341–360.
- [22] Mukhtar, B., Shad, M. K., Woon, L. F., Haider, M., & Waqas, A. (2024). Integrating ESG disclosure into the relationship between CSR and green organizational culture toward green innovation. *Social Responsibility Journal*, 20(2), 288–304. <https://doi.org/10.1108/SRJ-03-2022-0112>
- [23] Oman Vision 2040. (2024). Oman Vision 2040 strategic framework. <https://www.oman2040.om>
- [24] Plake, B. S., & Wise, S. L. (2014). Standards for educational and psychological testing. American Educational Research Association.
- [25] Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? *Research in Nursing & Health*, 29(5), 489–497.
- [26] Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? *Research in Nursing & Health*, 30(4), 459–467.
- [27] Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity. *Social Work Research*, 27(2), 94–104.
- [28] Suez. (2024). Suez sustainability and investment report 2024.
- [29] Times News Service. (2022). Oman Vision 2040 and sustainability transition. *Times of Oman*.
- [30] Waste Dive. (2024). Veolia sustainability performance and emissions targets. <https://www.wastedive.com>
- [31] World Bank. (2025). Oman economic diversification and sustainability report.
- [32] Zhang, J., Yu, C. H., Zhao, J., & Lee, C. C. (2025). How does corporate digital transformation affect green innovation? Evidence from China's enterprise data. *Energy Economics*, 142, 108217. <https://doi.org/10.1016/j.eneco.2025.108217>