

DEVELOPMENT, TRANSLATION AND PRELIMINARY VALIDATION OF DIGITAL TRANSFORMATION AND GREEN LOGISTICS MEASUREMENT INSTRUMENT

AL-HADIDI, O. K.^{1*} – SALLEH, N. A. M.¹

¹ *Faculty of Business and Accountancy, Universiti Selangor, Selangor, Malaysia.*

**Corresponding author
e-mail: 9232000011[at]student.unisel.edu.my*

(Received 11th August 2026; revised 09th October 2026; accepted 18th October 2026)

Abstract. This study develops, translates, and preliminarily validates a measurement instrument designed to assess digital transformation-driven green logistics practices and sustainability within cross-cultural logistics environments. A multi-phase methodological framework was applied, beginning with the conceptualization of constructs and the systematic generation of items based on established literature. Content and face validity were confirmed through expert evaluations, yielding strong CVI scores and demonstrating item relevance, clarity, and representativeness. The instrument was translated using standardized forward-backward procedures, followed by expert committee review to ensure semantic and conceptual equivalence. Cognitive debriefing further verified respondent comprehension, while a pilot test involving 30 logistics professionals provided evidence of acceptable internal consistency and survey feasibility. Structural validation procedures, such as Exploratory Factor Analysis (EFA), were intentionally deferred due to the small pilot sample size. Although the present study produced an Arabic version, the development and adaptation processes followed internationally recognized standards, making the instrument suitable for application in other linguistic and cultural contexts with appropriate adaptation. Overall, the study delivers a culturally attuned and conceptually robust pre-final instrument that lays the groundwork for future large-scale psychometric validation and empirical investigations into the role of digital transformation in advancing sustainable logistics practices.

Keywords: *digital transformation, green logistics, sustainability, measurement instrument, validation*

Introduction

Digital transformation has become a pivotal force reshaping the global logistics industry. The widespread integration of advanced technologies, such as artificial intelligence (AI), the Internet of Things (IoT), and big data analytics; has enabled logistics organizations to streamline supply chain operations, optimize transportation management, minimize waste, and enhance planning accuracy (Hilali et al., 2020). In an increasingly competitive global market, digital transformation is no longer optional; it represents a strategic imperative for organizations seeking operational excellence and long-term sustainability (Onyshchenko and Yudenko, 2019). Growing empirical evidence demonstrates that digitalization significantly improves logistics performance and facilitates the adoption of sustainable practices. Digital solutions reduce inefficiencies, enhance resource utilization, and support firms in transitioning toward environmentally responsible logistics systems (Verbivska et al., 2023; Yang, 2023; Yıldız and Çiğdem, 2022). By modernizing their digital infrastructure, companies create the conditions for green logistics practices, once challenging to implement; to become feasible, efficient, and economically viable (Xyugan and Lysochenko, 2023; Sun et al., 2021).

Green logistics practices, including emission reduction, sustainable transportation alternatives, optimized routing, energy conservation, and resource-efficient operations,

remain central to addressing the environmental impact of traditional logistics systems (Nagy and Szentesi, 2024; Umar et al., 2021). Digital transformation serves as a key enabler of these practices by enabling accurate data collection, real-time monitoring, and advanced analytics that support data-driven decision-making (Duc et al., 2024; Fan et al., 2024; Issa et al., 2024; Liu et al., 2024). Through enhanced supply-chain visibility and predictive capabilities, firms can reduce their environmental footprint while maintaining cost-effectiveness and operational performance (Umar et al., 2021; Vienažindienė et al., 2021). Despite the growing body of literature linking digital transformation and green logistics, validated measurement instruments in multilingual, cross-cultural contexts remain limited. Many existing tools were initially developed in non-Arabic environments and subsequently translated with varying levels of rigor. Such translations often lack systematic cultural adaptation and psychometric evaluation, increasing the risk of semantic bias and reduced construct validity when applied across cultural boundaries (Albahrain et al., 2024; Almohaya et al., 2024; Kanaan et al., 2022; Ehab et al., 2021). Prior research has emphasized that scientifically robust cross-cultural adaptation, supported by multi-stage translation and expert review, substantially enhances the accuracy, clarity, and conceptual equivalence of measurement instruments (Aldrees et al., 2018; Aziz et al., 2018).

Translation in cross-cultural research extends far beyond linguistic conversion; it is a conceptual and cognitive process aimed at preserving meaning, nuance, and contextual relevance across diverse cultural settings (Fieux et al., 2023; Almutary and Tayyib, 2021; Belayneh et al., 2019). Instruments that lack such rigor may fail to capture the intended constructs, resulting in measurement errors and compromised validity (Abrahamsen et al., 2023; Amri et al., 2022; Gonçalves et al., 2007). Accordingly, given the scarcity of rigorously adapted instruments to assess the nexus among digital transformation, green logistics practices, and sustainability across diverse cultural environments, the need for a methodologically sound measurement tool remains pronounced. While the present study implements its translation and adaptation procedures within an Arabic-language context, the methodological framework follows internationally recognized cross-cultural adaptation standards, ensuring broad applicability and transferability to diverse logistics environments.

Therefore, this study aims to develop, translate, and preliminarily validate a multidimensional questionnaire designed to measure three key constructs—digital transformation, green logistics practices, and sustainability. Through systematic item development, bilingual translation, expert-based content validation, and clarity assessment, this study provides a conceptually grounded and cross-culturally adaptable pre-final instrument. The resulting tool is suitable for both academic and applied research. It offers a robust foundation for future large-scale psychometric validation and international empirical investigations into how digital transformation drives sustainable green logistics practices. Translation in cross-cultural research extends far beyond linguistic conversion; it is a conceptual and cognitive process aimed at preserving meaning, nuance, and contextual relevance across diverse cultural settings (Fieux et al., 2023; Almutary and Tayyib, 2021; Belayneh et al., 2019). Instruments that lack such rigor may fail to capture the intended constructs, resulting in measurement errors and compromised validity (Abrahamsen et al., 2023; Amri et al., 2022; Gonçalves et al., 2007).

Accordingly, given the scarcity of rigorously adapted instruments to assess the nexus among digital transformation, green logistics practices, and sustainability across diverse

cultural environments, the need for a methodologically sound measurement tool remains pronounced. While the present study implements its translation and adaptation procedures within an Arabic-language context, the methodological framework follows internationally recognized cross-cultural adaptation standards, ensuring broad applicability and transferability to diverse logistics environments. Therefore, this study aims to develop, translate, and preliminarily validate a multidimensional questionnaire designed to measure three key constructs, digital transformation, green logistics practices, and sustainability. Through systematic item development, bilingual translation, expert-based content validation, and clarity assessment, this study provides a conceptually grounded and cross-culturally adaptable pre-final instrument. The resulting tool is suitable for both academic and applied research. It offers a robust foundation for future large-scale psychometric validation and international empirical investigations into how digital transformation drives sustainable green logistics practices.

Materials and Methods

This methodological study adopted a structured, multi-phase framework consistent with internationally recognized guidelines for instrument development, translation, and preliminary validation. The overall aim was to produce a conceptually grounded, linguistically accurate, and culturally adaptable measurement instrument capable of assessing digital transformation, green logistics practices, and sustainability. To ensure scientific rigor and cross-cultural applicability, the procedures followed established standards in scale development, covering construct conceptualization, item generation, and expert-based evaluation; alongside widely accepted cross-cultural adaptation practices that incorporate forward translation, reconciliation, back translation, and cognitive debriefing (Morait et al., 2019; Sousa and Rojjanasrirat, 2011). The methodological workflow was implemented across four integrated phases: (1) conceptualization and item development, (2) content and face validity assessment, (3) translation and linguistic equivalence procedures, and (4) cognitive debriefing, pilot testing, and preliminary finalization. Together, these phases ensured that the emerging instrument possessed strong conceptual foundations, linguistic clarity, and readiness for subsequent large-scale psychometric evaluation.

The study utilized an instrument-development and preliminary validation design aimed at producing a scientifically rigorous and culturally appropriate questionnaire. This design was selected to address the absence of standardized, methodologically validated tools for assessing digital transformation-driven green logistics and sustainability in Arabic-speaking contexts while maintaining applicability across diverse cultural environments. To achieve this, the research process adhered to internationally recognized guidelines for scale development and cross-cultural adaptation, ensuring conceptual accuracy, linguistic equivalence, and contextual relevance (Kalfoss, 2019; Lau et al., 2018; Sousa and Rojjanasrirat, 2011). Four interconnected phases structured the overall workflow. The first phase involved conceptualizing core constructs and developing items systematically based on an extensive review of validated models and empirical literature. The second phase focused on expert-based content and face validation to determine item relevance, clarity, and representativeness. The third phase comprised a structured translation and cross-cultural adaptation process aimed at achieving semantic and conceptual equivalence between the English and Arabic

versions. Finally, the fourth phase included cognitive debriefing and pilot testing to refine instrument clarity and assess preliminary feasibility before large-scale psychometric validation. Collectively, these procedures ensured that the resulting instrument was theoretically coherent, empirically grounded, and linguistically suitable for future research applications.

Phase 1: Conceptualization and item development

Phase one established the conceptual and theoretical foundation of the measurement instrument before any linguistic or cultural adaptation. Following internationally recognized guidelines for scale and questionnaire development, this phase ensured that the emerging tool was theoretically sound, conceptually coherent, and suitable for subsequent validation (Lau et al., 2018; Tsang et al., 2017; Lietz, 2010). The process began with an extensive literature review covering digital transformation, green logistics practices, and sustainability, drawing on validated models and empirical studies from logistics, operations management, ICT transformation, and sustainability research. This review enabled clear construct definition, domain mapping, and delineation of the instrument's conceptual scope, consistent with methodological recommendations for evidence-based item development in health, education, and behavioral measurement research. Based on the established conceptual framework, an initial pool of items was systematically generated using insights from prior validated instruments and recurring thematic patterns in the literature. Item generation emphasized clarity, unidimensionality, representativeness, and alignment with the predefined constructs, reflecting best practices highlighted in bilingual and cross-cultural questionnaire development studies. The preliminary item pool was then reviewed internally, with the research team examining redundancy, conceptual overlap, and structural coherence. Items were refined, merged, or removed to ensure proportional representation of construct domains and strengthen alignment with the theoretical framework. This internal screening process mirrors pre-validation refinement procedures commonly recommended in methodological guidelines for instrument development. A visual summary of the sequential workflow implemented in Phase One, including literature review, construct definition, theoretical framework development, item generation, and internal refinement, is presented in *Figure 1*, illustrating how these steps collectively produced the first draft of the measurement instrument.

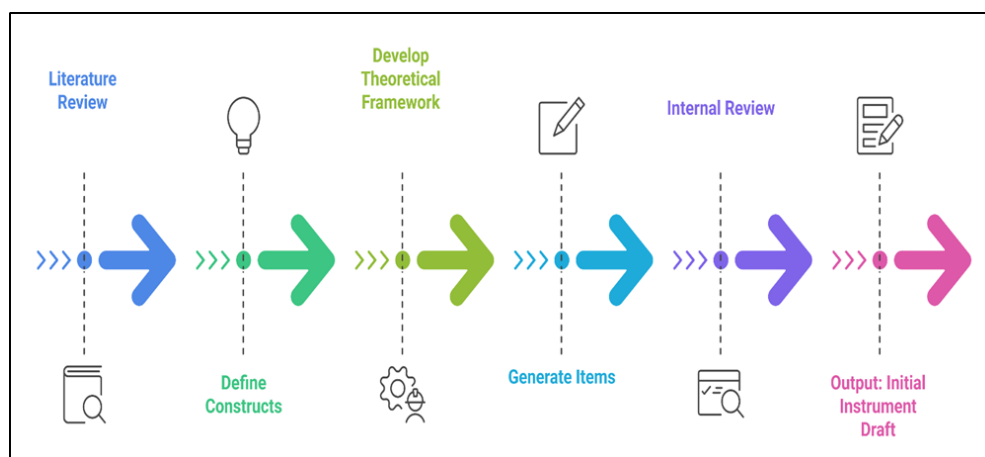


Figure 1. Workflow of Phase one: Conceptual and scientific development of the measurement instrument.

Phase two: Content and face validity assessment

Phase Two focused on the expert-based empirical refinement of the preliminary item pool to ensure conceptual adequacy, representativeness, clarity, and contextual relevance. This phase followed established standards in early-stage validation and reflects widely documented practices in instrument development and cross-cultural research (Makki et al., 2021; Yusoff, 2019; Lau et al., 2018). Through structured content and face validity procedures, the items generated in Phase One were evaluated and refined to ensure their suitability for assessing digital transformation, green logistics practices, and sustainability within logistics environments.

Content validity assessment

Content validity was conducted through a panel of academic experts representing digital transformation, logistics, sustainability, and measurement methodology. Using the Content Validity Index (CVI), experts independently evaluated each item for relevance, clarity, and representativeness (Yusoff, 2019). Item-level (I-CVI) and scale-level (S-CVI/Ave) values were calculated to quantify the adequacy of the item set, following procedures widely applied in validated adaptation studies. Items falling below recommended thresholds were reconsidered based on detailed expert comments (Park et al., 2017; Polit and Beck, 2006). Qualitative feedback highlighted conceptual overlap, ambiguous wording, and underrepresentation of specific sustainability dimensions, guiding revisions that strengthened structural clarity and alignment with the theoretical framework established in Phase one.

Face validity assessment

Following expert refinement, the revised version underwent face validity assessment by logistics and sustainability practitioners to ensure readability, interpretability, and operational applicability. Practitioners reviewed item wording within the context of real-world logistics workflows, providing feedback that informed further linguistic simplification, clarification of technical expressions, and adjustment of context-specific terminology. This stage ensured that the instrument would be consistently interpreted across respondents with diverse operational backgrounds, in accordance with recommendations found in questionnaire design and adaptation literature (Makki et al., 2021; Lau et al., 2018). The outcome of Phase Two was a fully refined English-language version of the instrument demonstrating strong content alignment, conceptual clarity, and practical comprehensibility. The integration of academic expert evaluation with practitioner review produced a theoretically robust and operationally relevant item set, forming a stable foundation for the translation and cross-cultural adaptation procedures undertaken in Phase Three. A visual summary of the Phase Two workflow, including expert appraisal, practitioner feedback, and item refinement, is shown in *Figure 2*.

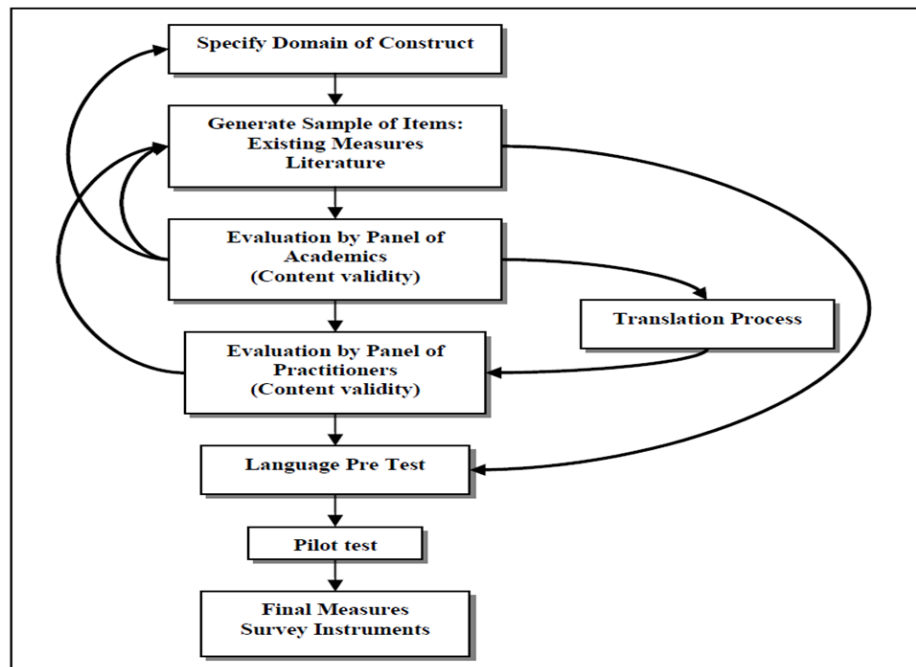


Figure 2. Integrated workflow for instrument development, validation, translation, and pretesting.

Source: Salleh (2006); Malhotra and Grover (1998).

Phase three: Translation and cross-cultural adaptation

Phase Three involved a structured translation and cross-cultural adaptation process to ensure that the Arabic version of the instrument achieved semantic, idiomatic, experiential, and conceptual equivalence with the original English version. The procedures followed internationally recognized guidelines for cross-cultural adaptation of self-report instruments and were informed by established methodological literature on questionnaire translation and test adaptation (Fenn et al., 2020; Sousa and Rojjanasrirat, 2011).

Forward translation

The translation began with two independent forward translations of the English instrument into Arabic. One translation was conducted by a professional linguist specializing in English–Arabic equivalence. At the same time, the second was performed by a subject-matter expert familiar with logistics, sustainability, and digital transformation terminology. This dual-translation strategy aligns with recommended cross-cultural adaptation practices, which advocate using multiple translators to ensure both linguistic naturalness and conceptual precision (Makki et al., 2021; Fenn et al., 2020).

Reconciliation

The two forward translations were then compared in a structured reconciliation process involving the translators and the research team. Discrepancies in terminology, conceptual emphasis, and contextual phrasing were discussed and resolved to produce a unified Arabic draft. This step reduces semantic drift and helps maintain fidelity to the

intended constructs, in line with best-practice recommendations in questionnaire and test design research (Lietz, 2010).

Back translation and expert committee review

The reconciled Arabic version was independently back-translated into English by bilingual translators who were not exposed to the original instrument. Differences between the back-translated and source versions were examined to identify deviations in meaning or conceptual representation, leading to iterative refinements. Back translation is widely recognized as a key quality-control step in cross-cultural instrument adaptation to safeguard conceptual equivalence across languages (Fenn et al., 2020). An expert committee comprising linguists, psychometricians, and domain specialists subsequently reviewed the Arabic version to evaluate semantic, idiomatic, experiential, and conceptual equivalence. Their feedback led to targeted revisions that enhanced clarity, cultural relevance, and consistency with the original constructs (Yusoff, 2019). The overall workflow of Phase Three, including forward translation, reconciliation, back translation, and expert committee review; is summarized in *Figure 3*, which visually depicts the sequential steps undertaken to ensure linguistic accuracy and conceptual integrity before pretesting.

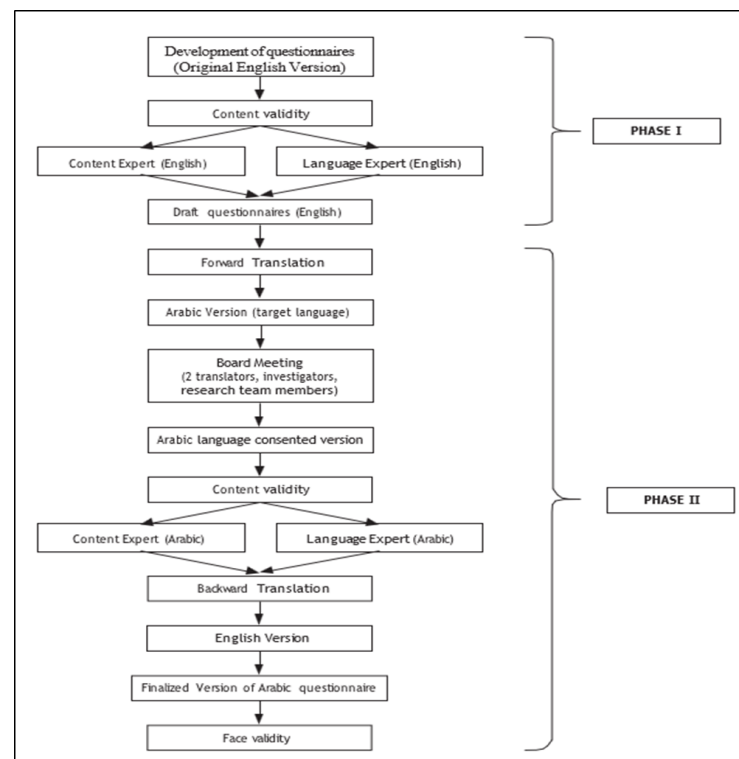


Figure 3. Multi-step workflow for forward translation, expert reconciliation, back translation, and bilingual content and face validation during the Arabic adaptation of the instrument.

Phase four: Cognitive debriefing, pilot testing, and preliminary finalization

Phase four aimed to ensure that the translated instrument was conceptually straightforward, culturally appropriate, and operationally feasible before undertaking large-scale psychometric validation. This phase combined qualitative cognitive

debriefing with quantitative pilot testing in accordance with internationally recognized guidelines for instrument refinement and early-stage measurement assessment (Ramu et al., 2023; Park et al., 2017; Polit and Beck, 2006). These procedures enabled the identification of interpretive issues, evaluation of item performance, and refinement of the pre-final instrument.

Cognitive debriefing

Cognitive debriefing was conducted with a purposive sample of approximately 5–10 participants, consistent with methodological recommendations emphasizing small, focused samples for in-depth assessment of item comprehension (Makki et al., 2021; Tsang et al., 2017). Participants were asked to paraphrase items, reflect on their interpretations, and identify unclear or culturally incongruent wording. Think-aloud and probing techniques were used to uncover discrepancies between intended and perceived meaning (Lietz, 2010). Insights obtained during this stage informed targeted revisions to enhance semantic clarity, improve contextual alignment, and refine terminology that could affect interpretation consistency across respondents.

Pilot testing

Following revisions derived from cognitive debriefing, a pilot test involving 20–30 logistics professionals was conducted, an empirically supported range for a preliminary field evaluation before complete psychometric testing (Ramu et al., 2023; Walde and Völlm, 2023). Participants completed the full instrument under conditions similar to the anticipated study environment. The pilot phase assessed item clarity, completion time, response variability, internal consistency trends, and the identification of items displaying ambiguity or conceptual weakness. In line with established pilot-study guidelines (Park et al., 2017; Polit and Beck, 2006), the purpose of the pilot test was not full psychometric validation but rather to detect early refinement needs. Minor adjustments were made to items exhibiting unclear phrasing or limited response variation.

Preliminary finalization

Findings from content and face validity, translation stages, cognitive debriefing, and pilot testing were consolidated to produce a pre-final version of the instrument. This integrative step ensured that the instrument met preliminary standards of clarity, reliability, and contextual adequacy (Hamed, 2022; Makki et al., 2021). Revisions were guided by conceptual coherence and linguistic precision, with redundant or weak items removed. A systematic revision log was maintained to support methodological transparency, consistent with best practices in instrument development and adaptation (Fenn et al., 2020). Although a complete psychometric evaluation was beyond the scope of this phase, the refined instrument is now prepared for subsequent large-scale validation, including exploratory and confirmatory factor analyses and assessments of convergent, discriminant, and criterion validity (Yusoff, 2019). The workflow of Phase Four, covering cognitive debriefing, iterative refinement, pilot testing, and preliminary finalization; is summarized in *Figure 4*.

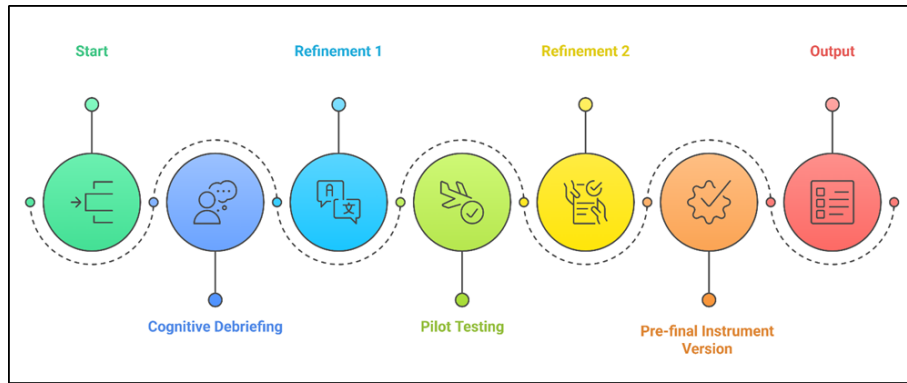


Figure 4. Workflow for Phase 4: Cognitive debriefing, pilot testing, and preliminary finalization.

Results and Discussion

This chapter presents the preliminary findings derived from the multi-stage development, validation, and adaptation procedures undertaken in this study. These early results provide empirical and qualitative evidence to evaluate the clarity, relevance, linguistic accuracy, and cultural appropriateness of the newly developed measurement instrument before conducting complete psychometric testing. Early-stage validation is widely recognized as an essential step in establishing foundational reliability and content adequacy before large-scale data collection and structural validation (Yusoff, 2019; Park et al., 2017; Polit and Beck, 2006). The chapter first reports the outcomes of the content validity assessment conducted by academic experts, including item-level and scale-level Content Validity Index (CVI) values, which indicate the conceptual representativeness and relevance of each item (Fenn et al., 2020; Yusoff, 2019). This is followed by findings from the face validity review conducted by logistics and sustainability practitioners, which offer insights into item clarity, interpretability, and practical applicability in operational contexts (Lietz, 2010).

Subsequent sections summarize the outcomes of the translation and linguistic equivalence procedures: forward translation, reconciliation, back translation, and expert committee review; that ensured semantic and conceptual consistency between the English and Arabic versions (Makki et al., 2021; Sousa and Rojjanasrirat, 2011). The results of cognitive debriefing sessions are then presented, highlighting comprehension issues, cultural nuances, and interpretive inconsistencies that informed further refinement of the instrument. Finally, the chapter reports the preliminary pilot test findings, including item performance, response distribution, completion time, and initial internal consistency patterns, key indicators of readiness for full-scale psychometric validation (Khanal and Chhetri, 2024; Walde and Völlm, 2023; Polit and Beck, 2006). Collectively, these preliminary results indicate that the instrument demonstrates satisfactory levels of conceptual clarity, linguistic accuracy, and functional usability, supporting its progression to the subsequent phase of comprehensive psychometric evaluation.

Content and face validity

Content and face validity constitute the foundational stages of preliminary instrument evaluation, providing complementary perspectives on the adequacy of measurement

items. Content validity examines the theoretical relevance, representativeness, and conceptual sufficiency of each item as assessed by academic experts, ensuring that the instrument accurately reflects the domains of digital transformation, green logistics practices, and sustainability. In contrast, face validity evaluates the readability, clarity, and contextual appropriateness of the instrument from the viewpoint of practitioners and potential respondents, confirming that items are understandable, operationally meaningful, and free from ambiguity (Yusoff, 2019; Polit and Beck, 2006). To guide the content validation process, the study employed the Content Validity Index (CVI) framework, which includes both item-level indices (I-CVI) and scale-level indices such as S-CVI/Ave and S-CVI/UA. These indices quantify expert agreement on item relevance using standardized rating criteria and are widely endorsed in methodological literature for early-stage instrument evaluation. As illustrated in *Figure 5*, the CVI procedures enabled the systematic identification of items requiring refinement or elimination. Face validity assessment complemented this process by capturing practitioner feedback related to wording clarity, interpretive simplicity, and contextual fit within logistics and sustainability environments. Together, content and face validity provided strong preliminary evidence of conceptual accuracy and practical usability, forming the basis for subsequent translation, cognitive debriefing, and pilot testing.

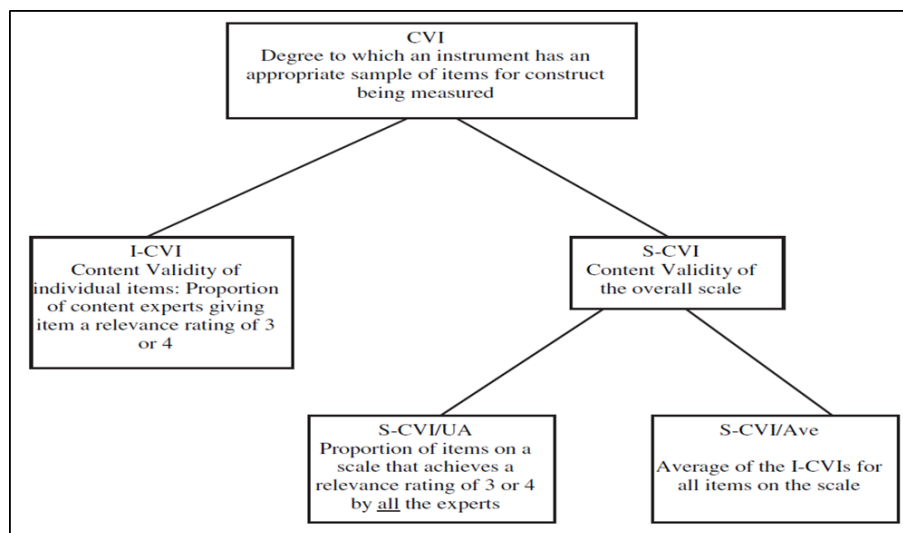


Figure 5. Structure of the Content Validity Index (CVI) framework used to assess item relevance and scale adequacy.

Content validity results

Content validity was assessed by eight academic experts who evaluated all items using a seven-point Likert scale ranging from 1 (“weak rating”) to 7 (“strong rating”). Consistent with widely accepted methodological guidelines, ratings of 5–7 were classified as indicative of item relevance, whereas ratings below 5 suggested the need for item modification (Yusoff, 2019; Polit and Beck, 2006). The Item-Level Content Validity Index (I-CVI) was calculated as the proportion of experts who assigned relevance scores. At the same time, scale-level indices (S-CVI/UA and S-CVI/Ave) were computed according to established procedures (*Table 1*). Table As shown in *Table 2. 1*, the I-CVI values for the 13 items ranged from 0.50 to 1.00, indicating strong overall agreement among experts regarding item relevance. Most items met or exceeded

the recommended threshold of 0.78, confirming adequate conceptual representation across the instrument. The mean I-CVI was 0.93, indicating excellent item-level content adequacy. At the scale level, the instrument showed similarly strong performance, with S-CVI/UA = 0.85 and an average expert-relevance proportion of 0.94, both exceeding the benchmarks commonly recommended for newly developed measurement instruments. Two items: Item 5 (I-CVI = 0.63) and Item 7 (I-CVI = 0.50); fell below the suggested cutoff and were therefore flagged for refinement. These items were subsequently reviewed and adjusted during later stages of instrument adaptation to enhance conceptual clarity and domain representativeness. Overall, the findings demonstrate that the instrument possesses strong preliminary content validity, with the majority of items judged by experts as clear, relevant, and theoretically aligned with their intended constructs.

Table 1. Content validity ratings, agreement scores, and I-CVI values from academic experts.

Item	A1	A2	A3	A4	A5	A6	A7	A8	NIA	I-CVI
Item 1	6	6	6	6	6	7	6	7	8	1
Item 2	6	7	6	6	6	6	6	6	8	1
Item 3	6	6	6	6	6	6	6	6	8	1
Item 4	6	7	6	7	6	6	6	6	8	1
Item 5	3	6	5	5	6	4	3	5	5	0.63
Item 6	6	7	6	6	6	6	7	7	8	1
Item 7	3	6	3	5	6	5	3	5	4	0.5
Item 8	6	6	6	7	6	6	7	6	8	1
Item 9	7	6	6	6	7	7	7	6	8	1
Item 10	7	6	6	7	7	6	6	7	8	1
Item 11	7	6	6	5	6	6	6	6	8	1
Item 12	7	7	6	7	6	6	7	6	8	1
Item 13	6	7	6	6	6	6	5	6	8	1
Mean I-CVI	-	-	-	-	-	-	-	-	-	0.93
S-CVI/UA	-	-	-	-	-	-	-	-	-	0.85
Proportion relevant	0.85	1	0.92	1	1	0.92	0.85	1	-	-
Academic proportion	-	-	-	-	-	-	-	-	-	0.94

Note: A1=Academic 1 until A8=Academic 8; NIA=Number In Agreement; I-CVI=Item-Level Content Validity Index; S-CVI/UA = Universal Agreement. Items are anonymized (Item 1, Item 2, ...) to protect instrument confidentiality prior to full validation. The agreement reflects the number of experts rating each item as relevant (scores of 5–7).

Table 2. Preliminary reliability (Cronbach's Alpha) for each construct in pilot test.

Construct	No. of items	Cronbach's α
Construct 1	4	.890
Construct 2	4	.941
Construct 3	4	.834
Construct 4	4	.896
Construct 5	5	.734
Construct 6	5	.916
Construct 7	5	.921
Construct 8	5	.734
Construct 9	6	.922
Construct 10	4	.912
Construct 11	3	.856
Construct 12	4	.881
Construct 13	5	.897

Note: Note: $\alpha \geq 0.70$ is considered acceptable for preliminary reliability.

Source: Polit and Beck (2006)

Face validity results

Face validity was assessed by a panel of logistics and sustainability practitioners to evaluate the instrument's clarity, readability, and ease of completion. Practitioners rated each item on a seven-point scale ranging from 1 ("difficult to understand") to 7 ("straightforward to understand"), following recommended procedures for evaluating surface-level comprehensibility in newly developed instruments (Lietz, 2010; Polit and Beck, 2006). Consistent with accepted criteria, ratings of 5–7 were interpreted as indicators of adequate clarity, whereas ratings below 5 suggested the need for potential refinement (Table 3). The Face Validity Index (I-CVI) was calculated for each item as the proportion of practitioners who assigned ratings within the acceptable clarity range. As shown in Table 3.2, I-CVI values ranged from 0.86 to 1.00, indicating consistently high levels of agreement among practitioners regarding the instrument's clarity and usability. The mean I-CVI reached 0.98, demonstrating excellent item-level face validity, while the scale-level index (S-CVI/UA = 0.85) exceeded commonly recommended thresholds for instruments in early development stages. Only a few items received comments suggesting simpler wording or more precise phrasing to enhance accessibility for respondents with diverse operational backgrounds. These refinements were incorporated during the subsequent revision phase. Overall, the findings confirm that the instrument exhibits strong preliminary face validity, with practitioners consistently rating the items as clear, easy to interpret, and suitable for logistics and sustainability contexts.

Table 3. Face validity ratings, agreement scores, and I-CVI values from industry professionals.

Item	E1	E2	E3	E4	E5	E6	E7	NIA	I-CVI
Item 1	7	6	7	6	6	7	6	7	1
Item 2	6	6	7	7	7	7	6	7	1
Item 3	7	7	6	7	6	6	7	7	1
Item 4	6	7	6	6	6	6	6	7	1
Item 5	6	6	4	5	6	5	5	6	0.86
Item 6	6	6	6	6	6	7	6	7	1
Item 7	6	6	6	6	6	5	6	7	1
Item 8	7	6	7	7	7	7	6	7	1
Item 9	6	7	6	6	6	6	7	7	1
Item 10	7	6	7	7	7	7	6	7	1
Item 11	6	7	6	6	6	6	6	7	1
Item 12	7	7	4	7	6	7	6	6	0.86
Item 13	7	6	6	7	6	7	6	7	1
Mean I-CVI	-	-	-	-	-	-	-	-	0.98
S-CVI/UA	-	-	-	-	-	-	-	-	0.85
Proportion relevant	1	1	0.85	1	1	1	1	-	-
Academic proportion	-	-	-	-	-	-	-	-	0.98

Note: A1=Academic 1 until A8=Academic 8; NIA=Number In Agreement; I-CVI=Item-Level Content Validity Index; S-CVI/UA = Universal Agreement.

Translation validation results

The translation validation procedures ensured that the Arabic version of the instrument maintained linguistic accuracy, conceptual fidelity, and contextual appropriateness relative to the original English version. A multi-stage cross-cultural adaptation protocol was applied in accordance with internationally recognized guidelines for translating self-report instruments (Makki et al., 2021; Sousa and Rojjanasrirat, 2011). The process began with two independent forward translations, produced separately by a bilingual linguistic expert and a subject-matter specialist. The

two translations were compared and merged during a structured reconciliation meeting, resulting in a unified version that demonstrated strong semantic accuracy and contextual fit. This reconciled draft was then subjected to two independent back-translations by bilingual translators who were blinded to the original English instrument. Comparing the back-translated and source versions enabled the identification and correction of subtle idiomatic inconsistencies, conceptual drift, and differences in interpretive nuance. Following the back-translation stage, an expert committee, comprising linguistic specialists, logistics practitioners, and methodological experts; conducted a comprehensive evaluation across the four established equivalence dimensions. Semantic equivalence ensured accurate transmission of meaning; idiomatic equivalence verified that expressions were rendered naturally rather than literally; experiential equivalence confirmed that references and examples were relevant to Arabic-speaking logistics contexts; and conceptual equivalence guaranteed that each item retained the theoretical intent of the original construct. This multi-perspective review strengthened the conceptual and linguistic integrity of the translated instrument. As an optional supplementary check, a small bilingual sample completed both language versions of the instrument to explore potential discrepancies in item interpretation. While not a required component of cross-cultural adaptation protocols, this step provided additional assurance that there was no meaningful difference between the English and Arabic versions. Finally, the Arabic version underwent a linguistic pre-test to assess clarity, fluency, and overall comprehensibility. Participants reported that the items were clear and culturally appropriate, and only minor modifications were required. Collectively, the translation validation procedures demonstrated strong semantic, idiomatic, experiential, and conceptual alignment between the two versions, confirming that the Arabic instrument was ready for subsequent cognitive debriefing and pilot testing.

Cognitive debriefing results

Cognitive debriefing was conducted to verify that the translated items were understood as intended and to identify any remaining issues related to clarity, interpretation, or cultural relevance. Using established cognitive interviewing techniques, including paraphrasing, probing, and think-aloud approaches; participants were asked to restate each item in their own words and comment on any difficulties encountered (Park et al., 2017; Polit and Beck, 2006). Overall, participants reported that the items were clear, meaningful, and easy to interpret. Only minor suggestions were offered, primarily related to simplifying a few technical expressions to enhance accessibility for respondents with varying levels of operational experience. These refinements were integrated into the revised version of the instrument. No concerns emerged regarding semantic distortion, cultural mismatch, or conceptual misunderstanding. The findings confirm that the translated instrument demonstrates satisfactory comprehensibility and cultural suitability, supporting its readiness for subsequent pilot testing.

Pilot testing results

A pilot test was administered to a sample of 30 logistics professionals to evaluate the instrument's clarity, usability, and preliminary reliability following completion of the translation, expert review, and cognitive debriefing stages. Feedback from participants indicated that the items were generally clear, contextually appropriate, and easy to

complete. No significant difficulties were reported, and the average completion time aligned with expectations for instruments of comparable scope, suggesting an acceptable respondent burden. Preliminary reliability was assessed using Cronbach's alpha for each construct. As presented in *Table 2*, all thirteen constructs demonstrated acceptable to excellent internal consistency, with alpha coefficients exceeding the standard benchmark of 0.70 recommended for early-stage reliability assessment (Polit and Beck, 2006). The obtained values ranged from 0.734 to 0.941, indicating that items within each construct performed cohesively and measured their intended domains consistently. Only minor refinements, primarily involving wording simplification; were implemented based on participant comments; no structural revisions or item deletions were required. The results confirm that the instrument performs reliably at the preliminary level and is suitable for progression to large-scale validation. Subsequent studies may therefore proceed with more advanced psychometric analyses, including exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and assessments of convergent and discriminant validity.

Finalization of the instrument

The cumulative findings from the content validity assessment, face validity review, translation validation procedures, cognitive debriefing, and pilot testing provided strong empirical and qualitative support for refining and finalizing the measurement instrument. Across these stages, all items demonstrated high levels of conceptual relevance, linguistic clarity, and cultural suitability, complemented by satisfactory preliminary reliability indices. Only minor wording refinements were required, and no substantive structural modifications were deemed necessary. Consequently, the instrument reached a stable, methodologically sound pre-final version, appropriate for advancing to a full-scale psychometric evaluation. An Exploratory Factor Analysis (EFA) was intentionally not conducted at this stage due to the limited pilot sample size, which was intended primarily for early refinement rather than structural validation. This approach is consistent with methodological recommendations stating that factor-analytic procedures require substantially larger samples, typically 150-300 participants or more; to produce stable factor solutions and meaningful construct interpretation (Park et al., 2017; Polit and Beck, 2006). Thus, the pilot test fulfilled its intended role of assessing clarity, usability, and preliminary internal consistency, without engaging in advanced statistical modeling that would be inappropriate and potentially misleading at this developmental phase. The results presented in this chapter confirm that the instrument is conceptually grounded, linguistically coherent, and operationally feasible, positioning it for subsequent large-scale validation. Future research will employ EFA, Confirmatory Factor Analysis (CFA), and additional reliability and validity assessments to establish the instrument's full psychometric properties and confirm its structural robustness.

Conceptual validation and practical applicability

This study developed, translated, and preliminarily validated a multidimensional instrument designed to assess digital transformation-driven green logistics practices and sustainability. By following a rigorous, multi-phase development process, the instrument achieved strong conceptual grounding, linguistic accuracy, and initial evidence of reliability, consistent with established standards for scale development and cross-cultural adaptation (Yusoff, 2019; Sousa and Rojjanasrirat, 2011). Content and

face validity results demonstrate that the instrument effectively represents the intended constructs and is suitable for use in applied logistics environments. High item-level and scale-level CVI values reflect strong expert agreement regarding conceptual relevance, aligning with methodological recommendations that emphasize expert judgment as a critical foundation for construct validity (Yusoff, 2019; Lietz, 2010). Items with lower CVI scores were refined, strengthening precision and construct alignment. Practitioner feedback further confirmed the instrument's clarity and operational relevance. Their recommendations primarily involved simplifying technical terminology to enhance usability—an expected pattern in applied domains such as logistics and sustainability. This aligns with established evidence showing that practitioner-informed revisions substantially improve surface-level clarity and contextual appropriateness (Lietz, 2010). Collectively, these results indicate that the instrument is both theoretically robust and practically suited for respondents working across different logistics functions.

Cross-cultural translation quality and comprehension

The structured translation and cross-cultural adaptation procedures produced an Arabic version that preserved the semantic, idiomatic, experiential, and conceptual properties of the original instrument. This is consistent with best-practice guidelines recommending multi-stage translation, reconciliation, back-translation, and expert committee review to ensure equivalence between language versions (Sousa and Rojjanasrirat, 2011). The optional bilingual comprehension checks further confirmed that no meaningful discrepancies in item interpretation emerged between English and Arabic respondents, supporting the linguistic stability of the translated items (Makki et al., 2021). Combined with expert review, these results verify that the Arabic instrument is culturally appropriate and suitable for use in logistics settings within Arabic-speaking contexts while retaining its conceptual integrity for potential cross-country comparisons.

Cognitive debriefing, pilot results, and readiness for full validation

Cognitive debriefing confirmed that respondents interpreted the items as intended, with only minor lexical refinements required. This aligns with the literature, which emphasizes cognitive interviewing as a crucial step for detecting subtle comprehension issues before large-scale testing (Polit and Beck, 2006). The pilot test provided further support for the instrument's clarity and functionality. Cronbach's alpha values across all constructs exceeded the accepted threshold of 0.70, indicating strong internal consistency for a preliminary validation stage (Walde and Völlm, 2023). Participants also reported no difficulties completing the instrument, reaffirming its feasibility and usability. Significantly, Exploratory Factor Analysis (EFA) was intentionally deferred due to the small pilot sample size, in accordance with methodological recommendations discouraging factor analysis on insufficient samples (Park et al., 2017; Polit and Beck, 2006). Structural validation will be conducted in future research using a larger, more representative sample to ensure stable and interpretable factor solutions.

Contribution and implications

This study offers one of the first systematically developed and culturally adapted instruments for examining digital transformation-driven green logistics and sustainability in the Arab region. The instrument provides a robust foundation for future research aimed at measuring organizations' digital readiness, adoption of green logistics

practices, and sustainability performance across diverse contexts. Beyond its regional relevance, the development and adaptation methodology aligns with international standards, ensuring that the instrument can be applied or adapted for use in non-Arabic settings. Its structured development framework and preliminary validation evidence make it suitable for future empirical studies employing EFA, CFA, and advanced reliability and validity testing.

Conclusion

This study developed, translated, and preliminarily validated a multidimensional measurement instrument intended to assess digital transformation-driven green logistics practices and sustainability. By employing a rigorous, internationally recognized multi-phase methodology, including conceptualization, expert-based content validation, systematic translation procedures, cognitive debriefing, and pilot testing; the instrument demonstrated strong conceptual relevance, linguistic clarity, cultural appropriateness, and satisfactory preliminary reliability. Although the present study produced an Arabic version of the instrument, the development and validation procedures followed globally established standards. As such, the instrument's methodological foundations, rather than its language, constitute its primary contribution. The tool can therefore be adapted for use in non-Arabic settings, requiring only linguistic and cultural tailoring in accordance with standard cross-cultural adaptation protocols. In this regard, the study provides a robust template that can support broader international applications examining the interplay between digital transformation, green logistics, and sustainability. Several limitations should be acknowledged. First, the validation procedures reported here were intentionally preliminary. The cognitive debriefing and pilot testing samples were small by design and suitable only for refinement, not structural validation. Accordingly, advanced psychometric analyses, such as Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), were not conducted, as methodological guidelines require substantially larger samples for these procedures. Second, although the pilot respondents were drawn from logistics environments, the instrument's applicability extends beyond this context. However, variation in digital transformation maturity or sustainability practices across sectors may influence item interpretation, warranting context-specific assessment in future studies. Third, although rigorous translation and cross-cultural processes were applied, minor linguistic or cultural nuances may emerge when adapting the instrument to other languages or regional contexts. Finally, criterion-related and predictive validity were not examined, as these analyses require larger and more diverse datasets. Future studies should conduct full-scale psychometric validation using larger and more diverse samples to establish the instrument's factor structure, reliability, and validity. Sequential application of EFA followed by CFA is necessary to confirm the constructs identified in the development phase. Further research should examine convergent, discriminant, and criterion-related validity to assess how the new instrument relates to existing measures of digital transformation, sustainability, and logistics performance. Cross-sector and cross-country studies are particularly encouraged to evaluate the instrument's broader applicability and to explore how contextual differences influence digital transformation-driven sustainability practices. Longitudinal research designs may also provide insights into how green logistics capabilities evolve as digital transformation progresses. Finally, adapting the instrument

into additional languages and cultural settings would extend its international relevance and support comparative studies across global logistics systems.

Acknowledgement

This research is self-funded.

Conflict of interest

The authors confirm that there is no conflict of interest involve with any parties in this research study.

REFERENCES

- [1] Abrahamsen, C., Jensen, C.M., Elkjær, M., Østervang, C., Knudsen, B.M., Jensen, S.S., Moos, C., Viberg, B., Mogensen, C.B. (2023): Translation and cultural adaption of the emergency department-consumer assessment of healthcare providers and systems (ED CAPHS)-a questionnaire to measure patient experience in Denmark. – *Scandinavian Journal of Caring Sciences* 37(3): 872-879.
- [2] Albabtain, B., Paudyal, V., Cheema, E., Bawazeer, G., Alqahtani, A., Bahatheq, A., Shuweihdi, F., Hadi, M.A. (2024): Translation, cultural adaptation and validation of Patient Satisfaction with Pharmacist Services Questionnaire (PSPSQ) 2.0 into the Arabic language among people with diabetes. – *PLOS ONE* 19(6): 13p.
- [3] Aldrees, T., Almubarak, Z., Hassounah, B., Albosaily, A., Aloulah, M., Almasoud, M., et al. (2018): Translation, validation, and cultural adaptation of the Rhinosinusitis Disability Index and the Chronic Sinusitis Survey into Arabic. – *Annals of Saudi Medicine* 38(3): 159-166.
- [4] Almohaya, H., Bakhsh, H., Sheeha, B., Aldhahi, M., Alhasani, R. (2024): Translation and cultural adaptation into Arabic of Patient-Reported Outcome Measurement Information System® item banks: Cognitive Function Abilities and Physical Function for samples with mobility aid users. – *Healthcare* 12(2): 10p.
- [5] Almutary, H., Tayyib, N. (2021): Translation, adaptation, and validation of chronic kidney disease self-management and knowledge instruments for people at pre-dialysis stage in the Arab world. – *Clinical Nursing Research* 30(8): 1231-1240.
- [6] Amri, M.I., Alzhirani, M.M., Alanazi, A.D., Alqahtani, M.M., Kashoo, F.Z. (2022): Cross-cultural adaptation and validation of the Arabic version of the Foot Function Index in patients with chronic lateral ankle instability. – *Journal of Foot and Ankle Research* 15(1): 11p.
- [7] Aziz, M., Galal, M., Elzohri, M., El-Nouby, F., Leong, K. (2018): Cross-cultural adaptation and validation of Systemic Lupus Erythematosus Quality of Life Questionnaire into Arabic. – *Lupus* 27(5): 780-787.
- [8] Belayneh, T., Gebeyehu, A., Adefris, M., Rørtveit, G., Genet, T. (2019): Translation, transcultural adaptation, reliability and validation of the Pelvic Organ Prolapse Quality of Life (P-QoL) in Amharic. – *Health and Quality of Life Outcomes* 17(1): 10p.
- [9] Duc, A., Nguyen, T., Nguyen, T., Nguyen, T., Duong, T. (2024): The role of infrastructure in green logistics development: A case study of Vietnamese logistics enterprises. – *European Journal of Business and Management Research* 9(3): 135-141.
- [10] Ehab, B., Hussein, R., Abdullah, E., Ahmed, Z., Elsherbiny, R. (2021): Cultural adaptation and validation of the EORTC QLQ-BR45 to assess health-related quality of life of breast cancer patients. – *European Pharmaceutical Journal* 68(2): 41-48.

- [11] Fan, Z., Zhi-qiang, Z., Wang, X., Sun, Z. (2024): The impact of digital economy on the development of green logistics under the goal of dual carbon. – *Journal of Education, Humanities and Social Sciences* 44: 19-30.
- [12] Fenn, J., Tan, C., George, S. (2020): Development, validation and translation of psychological tests. – *BJPsych Advances* 26(5): 306-315.
- [13] Fieux, M., Carsuzaa, F., Charriot, J., Michel, J., Subtil, F., Sowerby, L., et al. (2023): Cross-cultural adaptation of the CRS-PRO questionnaire into French. – *Journal of Otolaryngology - Head & Neck Surgery* 52(1): 10p.
- [14] Gonçalves, C., Peccin, M., Almeida, G., Cohen, M. (2007): Tradução, validação e adaptação cultural da escala de atividade esportiva. – *Acta Ortopédica Brasileira* 15(5): 246-250.
- [15] Hamed, T. (2022): Designing a questionnaire for a research paper: A comprehensive guide to design and develop an effective questionnaire. – *Asian Journal of Managerial Science* 11(1): 8-16.
- [16] Hilali, W., Manouar, A., Idrissi, M. (2020): Reaching sustainability during a digital transformation: A PLS approach. – *International Journal of Innovation Science* 12(1): 52-79.
- [17] Issa, A., Khadem, A., Alzubi, A., Berberoğlu, A. (2024): The path from green innovation to supply chain resilience: Do structural and dynamic supply chain complexity matter? – *Sustainability* 16(9): 33p.
- [18] Kalfoss, M. (2019): Translation and adaption of questionnaires: A nursing challenge. – *SAGE Open Nursing* 5: 1-13.
- [19] Kanaan, S., Al-Karasneh, M., Almhdawi, K., Oteir, A., Guimarães, I., Al-Omari, A., et al. (2022): Translation and cross-cultural adaptation of the Low Back Pain Treatment Belief Questionnaire and its clinimetric properties. – *Journal of Back and Musculoskeletal Rehabilitation* 35(6): 1345-1355.
- [20] Khanal, B., Chhetri, D.B. (2024): A pilot study approach to assessing the reliability and validity of relevancy and efficacy survey scale. – *Janabhawana Research Journal* 3(1): 35-49.
- [21] Lau, A., Yusoff, M., Lee, Y., Choi, S., Xiao, J., Liong, M. (2018): Development and validation of a Chinese translated questionnaire: A single simultaneous tool for assessing gastrointestinal and upper respiratory tract related illnesses in pre-school children. – *Journal of Taibah University Medical Sciences* 13(2): 135-141.
- [22] Lietz, P. (2010): Research into questionnaire design: A summary of the literature. – *International Journal of Market Research* 52(2): 249-272.
- [23] Liu, Z., Zhao, Y., Guo, C., Xin, Z. (2024): Research on the impact of digital-real integration on logistics industrial transformation and upgrading under green economy. – *Sustainability* 16(14): 18p.
- [24] Makki, M., Hassali, M., Awaisu, A., Chemaitelly, H., Hail, M., ElMotasim, W., et al. (2021): Development, translation, and validation of a bilingual questionnaire on unused medications in homes. – *Saudi Pharmaceutical Journal* 29(7): 648-655.
- [25] Malhotra, M.K., Grover, V. (1998): An assessment of survey research in POM: From constructs to theory. – *Journal of Operations Management* 16(4): 407-425.
- [26] Morait, S., Abuhaimed, M., Alharbi, M., Almohsen, B., Alturki, A., Alarbash, A. (2019): Attitudes and acceptance of the Saudi population toward cosmetic surgeries in Riyadh, Saudi Arabia. – *Journal of Family Medicine and Primary Care* 8(5): 1685-1690.
- [27] Nagy, G., Szentesi, S. (2024): Green logistics: Transforming supply chains for a sustainable future. – *Advanced Logistic Systems - Theory and Practice* 18(3): 29-42.
- [28] Onyshchenko, O., Yudenko, Y. (2019): Innovative activity of logistics enterprises in the modern world. – *SHS Web of Conferences* 67: 6p.
- [29] Park, C.S.-Y., Yoon, S.L., Yun, S.-N., Park, E. (2017): Korean Patient-Perceived Satisfaction Scale of Community-Based Case Management Services (Korean-PSCCM):

- Development and psychometric evaluation. – *Journal of Community Health Nursing* 34(1): 32-45.
- [30] Polit, D.F., Beck, C.T. (2006): The content validity index: Are you sure you know what's being reported? Critique and recommendations. – *Research in Nursing & Health* 29(5): 489-497.
- [31] Ramu, P., Osman, M., Abdul Mutalib, N.A., Aljaberi, M.A., Lee, K.-H., Lin, C.-Y., Hamat, R.A. (2023): Validity and reliability of a questionnaire on the knowledge, attitudes, perceptions and practices toward food poisoning among Malaysian secondary school students: A pilot study. – *Healthcare* 11(6): 14p.
- [32] Salleh, M. (2006): The effect of enacted capabilities on adoption and utilisation of innovative information systems: A study of small- and medium-sized enterprises. – The University of Queensland 250p.
- [33] Sousa, V.D., Rojjanasrirat, W. (2011): Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. – *Journal of Evaluation in Clinical Practice* 17(2): 268-274.
- [34] Sun, M., Grondys, K., Hajiyeve, N., Zhukov, P. (2021): Improving the e-commerce business model in a sustainable environment. – *Sustainability* 13(22): 23p.
- [35] Tsang, S., Royse, C.F., Terkawi, A.S. (2017): Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. – *Saudi Journal of Anaesthesia* 11(Suppl. 1): S80-S89.
- [36] Umar, M., Khan, S., Zia-ul-Haq, H., Yusliza, M., Farooq, K. (2021): The role of emerging technologies in implementing green practices to achieve sustainable operations. – *The TQM Journal* 34(2): 232-249.
- [37] Verbivska, L., Zhygalkevych, Z., Fisun, Y., Chobitok, I., Shvedkyi, V. (2023): Digital technologies as a tool of efficient logistics. – *Revista de la Universidad del Zulia* 14(39): 492-508.
- [38] Vienažindienė, M., Tamulienė, V., Zaleckienė, J. (2021): Green logistics practices seeking development of sustainability: Evidence from Lithuanian transportation and logistics companies. – *Energies* 14(22): 18p.
- [39] Walde, P., Völlm, B.A. (2023): The TRAPD approach as a method for questionnaire translation. – *Frontiers in Psychiatry* 14: 12p.
- [40] Xyugan, W., Lysochenko, A. (2023): Improving the supply chain management of China's rail logistics system. – *E3S Web of Conferences* 403: 8p.
- [41] Yang, C. (2023): Research on the decision-making mechanism of digital transformation of small and medium-sized logistics enterprises in China based on evolutionary game. – *Academic Journal of Management and Social Sciences* 5(2): 126-132.
- [42] Yıldız, B., Çiğdem, Ş. (2022): Digital transformation in logistics: Lights-out logistics. – In: *Financial and Economic Issues in Emerging Markets*, Özgür Publications 24p.
- [43] Yusoff, M.S.B. (2019): ABC of content validation and content validity index calculation. – *Education in Medicine Journal* 11(2): 49-54.