Idealogy Journal Vol. 11 No. 2, 2026, pp 828-849
DOI: https://doi.org/10.24191/idealogy.v11i2.1027
Section: Original Article

Expert User Validation of Perceived Quality and Purchase Intention: Perspectives from Product Designers and Industry Practitioners

***Mohamad Hasriq Haiqal Mohd Rashid[1], Haszlin Shaharudin[2], Nur Shahidatul Aina[3]**
[1,2,3] *Faculty of Art & Design, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia*

[1]haiqalhasriq@gmail.com[1], *haszlin@uitm.edu.my[2] , shieda.firzie66@gmail.com[3]
***Corresponding author**

Received: 11 August 2026; Accepted: 30 August 2026; Published: 1 September 2026

ABSTRACT

A design model derived from user evidence requires professional content validation before it is used as a structured reference for design decisions. This study evaluates the content validity of a Sustainable Product Design Model (SPDM) comprising 37 indicators across seven criteria: form, function, safety, sustainability, quality, commerciality, and innovation. Thirty product designers and industry practitioners, each with at least five years of professional experience, completed a four-point relevance questionnaire administered face to face. Item level Content Validity Index (I-CVI), Scale level Content Validity Index using the average method (S-CVI/Ave), and modified kappa were calculated. All 37 indicators met the predetermined I-CVI retention threshold of 0.78. I-CVI values ranged from 0.80 to 1.00, with an overall mean of 0.94, while criterion level S-CVI/Ave values ranged from 0.88 to 0.98. Function recorded the strongest consensus (0.98), followed by quality (0.97) and form (0.95), whereas innovation recorded the widest variation in expert agreement (0.88). The analysis also examined indicators with lower agreement and the effect of converting the four-point scale into relevant and lower relevance categories. Stability, Price, Prestige Value, and Exclusiveness/Limited met the retention criterion but required clearer operational definitions. The study establishes a dual evidence validation approach that preserves priority derived from user evidence while expert evidence establishes content relevance, conceptual boundaries, and professional support for sustainable product design decisions.

Keywords: *Content validity, Expert validation, Perceived product quality, Sustainable product design, Design decision support*

ISSN: 2550-214X © 2026. Published for Idealogy Journal by UiTM Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution-No Commercial-No Derivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

1 INTRODUCTION

Perceived product quality extends beyond technical conformance. Users judge quality through product appearance, ease of use, reliability, durability, material cues, information, brand meaning, service expectations, and prior experience. Golder et al. (2012) describe quality through production, experience, and evaluation, while Stylidis et al. (2020) organise perceived quality into attributes that support design prioritisation. This distinction is important because technical adequacy alone does not ensure user acceptance when value, usability, desirability, or trust is not recognised.

Sustainable product design makes this relationship more demanding. Environmental performance alone does not ensure adoption, continued use, repair, or retention. Bayram et al. (2024) show that product quality in circular economy research covers market value, customer perspectives, functionality, technical attributes, longevity, design, and environmental aspects. Körkemeyer et al. (2025) further demonstrate that sustainability operates as a dimension of perceived quality. Recent design research also shows that sustainability cues interact with product category and user perception rather than

functioning as universal purchase signals (Sierra-Fontalvo et al., 2026). These findings shift the design problem from simply including sustainable features to integrating sustainability with the attributes through which users judge product value.

The Sustainable Product Design Model (SPDM) examined in this study addresses this integration challenge. Its operational structure contains seven criteria: FORM, FUNCTION, SAFETY, SUSTAINABILITY, QUALITY, COMMERCIALITY, and INNOVATION, with 37 indicators derived from user evidence. The model supports early design and concept evaluation, where decisions about form, usability, materials, performance, market positioning, and lifecycle implications remain open to development. The priority order within each criterion was established from earlier user evidence. The present study retains that priority structure and evaluates the relevance and conceptual appropriateness of the indicators through judgements from experienced product designers and industry practitioners.

The distinction between user priority and expert relevance is central to the method. User evidence identifies which attributes matter most to intended users. Expert content validation examines how well those attributes represent the conceptual domains of the framework. Expert relevance scores provide content evidence, while the user data retain their original role in prioritisation. Combining both evidence types into one ranking changes the empirical role of each group and weakens the logic of the model.

This study strengthens the expert validation stage through four analytical components. First, it applies item level and scale level content validity analysis to all 37 indicators. Second, it applies modified kappa to examine agreement after correction for chance. Third, it interprets indicators with lower consensus as evidence for conceptual refinement rather than using expert scores to reorder the priority derived from user evidence. Fourth, it defines the limits of content validity statistics, which do not establish construct separation, predictive performance, causal effects, or universal applicability. The contribution is a dual evidence architecture in which user priorities and professional content judgements remain analytically distinct and work together to strengthen the SPDM.

2 LITERATURE REVIEW

2.1 Perceived Product Quality, Purchase Intention and Sustainable Choice

Perceived product quality is the overall judgement users form about a product's excellence or superiority relative to expectations and alternatives. It differs from objective quality because users interpret product and market cues rather than relying only on engineering measurement. Agarwal and Teas (2004) connect perceived quality with perceived value, sacrifice, and risk, while Beneke et al. (2013) show that perceived quality contributes to willingness to buy through perceived value. These relationships position product quality as both a design input and a market related user outcome.

Purchase intention represents a behavioural predisposition rather than a completed transaction, but it remains useful because it shows how evaluations are translated into willingness to select a product (Morwitz, 2014). Tsiotsou (2006) links perceived quality directly with purchase intention and satisfaction. In sustainable product contexts, environmental and circular benefits are evaluated together with price, appearance, material choice, repairability, complexity, and performance. The design task is therefore to express sustainability as product value without weakening the quality signals on which users rely.

Recent sustainable product studies strengthen this interpretation. Visser et al. (2018) show that household appliance purchase decisions combine reliability, durability, features, brand, value, and sustainability considerations, while Marcon et al. (2022) identify relationships between green product attributes and consumer behaviour. Bayram et al. (2024) further show that circular economy research repeatedly links product quality with functionality, longevity, technical attributes, design, customer

value, and environmental aspects. Mohd Rashid et al. (2026) report that design and quality remain prominent purchase considerations and organise consumer evidence around the same seven broad design domains. Alias and Che Din (2025) also show that visible design elements such as colour and imagery influence purchase decisions. Together, these studies support a framework in which sustainable value is integrated with quality, usability, market, and design judgements.

2.2 Sustainable Product Design as Decision Support in the Early Design Stage

Sustainable product design requires environmental, social, and economic considerations while product requirements and concepts are still being developed. Bocken et al. (2016) identify product life extension, reuse, recovery, and resource efficiency as central strategies in circular design. Isaksson Hallstedt et al. (2023) similarly emphasise decision support that guides sustainability during early development. These perspectives show that sustainability becomes actionable when broad goals are translated into clear design criteria before architecture, materials, interfaces, and service strategies become difficult to change.

User centred design provides an essential perspective on adoption. Veryzer and Borja de Mozota (2005) show that user-oriented design strengthens new product development by integrating user knowledge into design decisions. Du and Ayob (2024) position sustainability as a product design concern rather than a separate environmental addition. Abdul Jamil and Abu Bakar (2025) demonstrate that design development contributes to public perception, brand image, and purchase intention. Md. Nazri et al. (2025) show that aesthetics, ergonomics, sensory stimulation, and interaction shape product experience, while Mohd Hasbullah et al. (2025) demonstrate the integration of human centred design with a multidimensional sustainability framework. Together, these studies support a design logic in which technical responsibility, user experience, social meaning, and market acceptance are considered together.

The SPDM operationalises this convergence through seven criteria. FORM captures visible and tactile cues; FUNCTION captures comprehension, ergonomics, comfort, usability, and performance; SAFETY captures standards, testing, maintenance, warranty, and parts support; SUSTAINABILITY captures material, energy, disassembly, and recyclability concerns across the product lifecycle; QUALITY captures reliability, durability, specifications, material suitability, and conformance; COMMERCIALITY captures market related information and support; and INNOVATION captures user benefit, differentiation, status related value, exclusivity, and novelty. These criteria connect technical, user, lifecycle, and market concerns. Professional validation is therefore required before the framework is used as a structured design reference.

2.3 Expert Content Validation and the Logic of Agreement

Content validation examines how well a set of items represents an intended conceptual domain. Lynn (1986) established a systematic approach for quantifying expert judgements, while Davis (1992) emphasised structured expert review during instrument development. Polit and Beck (2006) distinguish item level content validity from scale level aggregation and state that researchers should report clearly how scale scores are calculated. Polit et al. (2007) also recommend accounting for chance agreement when interpreting item level content validity. Almanasreh et al. (2019) emphasise that rigorous content validation requires explicit domain definition, evidence of representativeness, transparent expert judgement procedures, and revision after quantitative appraisal. A high CVI therefore requires clear construct boundaries to support meaningful interpretation.

The I-CVI is the proportion of experts who rate an item as relevant. A four point scale is commonly used because it removes a neutral midpoint and requires a directional relevance judgement. The S-CVI/Ave is the average of the I-CVI values across items within a scale or domain. Yusoff (2019) presents a practical sequence for systematic content validation, and Zamanzadeh et al. (2015) show why

the average approach remains informative when universal agreement becomes difficult to achieve with larger expert panels. Almanasreh et al. (2019) also state that quantitative indices form one part of content evidence and do not replace conceptual judgement. Modified kappa provides a correction for chance agreement. With large expert panels and high agreement, its value remains close to I-CVI, so it is used as a robustness check rather than as an independent form of validity.

For a design model, these statistics provide evidence about relevance and representativeness. Their scope is content validity, while predictive validity, construct validity, and market performance require separate forms of evidence. Expert validation examines the coherence of terminology and the conceptual fit of items within their proposed domains. Evidence on design outcomes and any revision of user priorities require separate empirical procedures.

2.4 Critical Boundary Conditions Across the Seven Design Criteria

The seven SPDM criteria differ in their direct connection with observable product properties. FORM, FUNCTION, and QUALITY contain attributes encountered through perception, use, performance, or conformance. This direct connection with observable product experience is consistent with the stronger expert convergence found for these criteria. Research on product experience also shows that visible, tactile, ergonomic, and functional cues influence quality interpretation. Md. Nazri et al. (2025) and Alias and Che Din (2025) demonstrate that sensory and visual design elements shape user response and purchase evaluation.

SAFETY and SUSTAINABILITY extend beyond immediate product performance into lifecycle support. Mesa et al. (2022) show that durability, repair, refurbishment, and remanufacture are central to product life extension, while Amend et al. (2022) demonstrate that modular design and repair support influence repair behaviour and user experience. Munten and Vanhamme (2023) further show that repairability communication operates as a quality signal through perceived durability. These findings support the inclusion of maintenance and spare part availability and also place these indicators across the boundaries of safety, quality, and sustainability. Their analytical role therefore requires a clear operational definition.

COMMERCIALITY and INNOVATION depend more strongly on market and symbolic meaning. Meta analytic evidence shows that electronic word of mouth, message usefulness, credibility, and related digital review mechanisms affect intention to buy (Ismagilova et al., 2020). Innovation research also demonstrates that objective design newness and consumers' subjective perception of newness are distinct (Sung et al., 2022). Wang et al. (2022) distinguish exclusivity from rarity within luxury perception. These distinctions explain the wider spread of expert judgement for Price, Prestige Value, and Exclusiveness/Limited because these indicators also operate as variables of market positioning and are not direct measures of technical innovation.

The literature provides a clear basis for interpreting the validation results. Lower agreement in market related or symbolic criteria does not indicate weak validity when the items meet the retention threshold. Instead, the pattern reflects stronger dependence on context and construct definition. A rigorous validation process therefore distinguishes irrelevant items from relevant items that require explicit boundary conditions.

2.5 Research Gap, Novelty and Study Positioning

The literature provides strong foundations for perceived quality, sustainable design, user centred development, market acceptance, and content validation, but these areas are often examined separately. Prior studies demonstrate the importance of user evaluation, design development, sustainability, and product experience (Abdul Jamil & Abu Bakar, 2025; Alias & Che Din, 2025; Du & Ayob, 2024; Md. Nazri et al., 2025; Mohd Rashid et al., 2026). A methodological gap remains in the professional validation of design priorities derived from user evidence while preserving the empirical role of that

evidence. This study addresses that gap. Its novelty lies in separating validation from reranking and in using variation in expert agreement to identify construct boundaries. Table 1 positions the study against the main areas of established knowledge.

Table 1 Research Gap and Study Positioning

Research area	Established knowledge	Gap addressed	Contribution of this study
Perceived product quality	Quality is formed through user evaluation of product and market cues (Golder et al., 2012; Stylidis et al., 2020).	Attributes derived from user evidence require professional confirmation that they represent coherent design domains.	Validates 37 derived from user evidence indicators against seven proposed design criteria.
Sustainable product design	Early design decisions shape lifecycle and circular outcomes (Bocken et al., 2016; Isaksson Hallstedt et al., 2023).	Sustainability criteria are often addressed separately from perceived quality and market related criteria.	Tests sustainability together with form, function, safety, quality, commerciality, and innovation.
Content validity	I-CVI and S-CVI/Ave quantifies expert relevance judgements (Polit & Beck, 2006; Yusoff, 2019).	Using validation results as rankings misrepresents relevance evidence.	Separates expert relevance from the existing user priority order.
Chance agreement	Modified kappa adjusts item agreement for chance (Polit et al., 2007; Zamanzadeh et al., 2015).	Validation of design frameworks often reports raw agreement only.	Applies modified kappa to all 37 indicators.
Conceptual refinement	Expert review identifies unclear or overlapping content.	Lower agreement is often treated as a deletion trigger without diagnosing conceptual boundaries.	Uses lower consensus items to refine definitions while retaining empirically supported structure.

(Source: Authors' synthesis)

2.6 Operational Definitions and Analytical Architecture

The study keeps user priority, expert relevance, and conceptual refinement analytically separate. This prevents one evidence source from being interpreted as if it measured another construct. The operational definitions in Table 2 establish the role of each evidence layer.

2.7 Literature Alignment of the Seven Design Criteria

The seven criteria cover sensory, functional, lifecycle, quality, and market domains. Table 3 links each criterion with its conceptual role, supporting literature, and validation focus.

Table 2 Operational Definitions and Analytical Architecture

Construct	Operational meaning	Measurement	Analytical role
Priority derived from user evidence	Relative importance assigned to indicators by intended users in prior empirical work.	Existing priority order within each of the seven criteria.	Preserved as the model's prioritisation layer; not recalculated from expert scores.
Expert relevance	Professional judgement that an indicator appropriately represents its proposed criterion.	Four-point relevance scale completed by 30 experts.	Quantified using I-CVI and S-CVI/Ave.
Agreement corrected for chance	Expert agreement adjusted for the probability of chance concurrence.	Modified kappa derived from I-CVI and panel size.	Tests robustness of item level relevance agreement.
Conceptual refinement	Need to sharpen wording, scope, or criterion boundaries.	Comparative pattern of lower I-CVI values plus theoretical interpretation.	Guides operational clarification without automatic deletion or reranking.
Validated SPDM architecture	Structure of seven criteria supported by user and expert evidence serving different functions.	Integration of prior user priority with current expert content validity.	Provides decision support logic for early sustainable product development.

(Source: Authors' analytical framing)

Table 3 Literature Alignment of the Seven Design Criteria

Criterion	Conceptual role	Literature support	Validation focus
Form	Visible and tactile product cues that shape first impressions, affective response, and perceived quality.	Creusen & Schoormans (2005); Md. Nazri et al. (2025); Alias & Che Din (2025).	Assessment of the coherence of shape, appearance, colour, finishing, and texture as indicators of formal quality, with sensory experience clearly distinguished from aesthetics alone.
Function	Interaction between users and products through comprehension, ergonomics, comfort, usability, and performance.	Golder et al. (2012); Bayram et al. (2024); Mohd Rashid et al. (2026).	Assessment of functional quality across technical performance, comprehension, ergonomics, comfort, and ease of use.
Safety	Direct safety evidence plus lifecycle support that enables continued safe and dependable operation.	Mesa et al. (2022); Amend et al. (2022); Munten & Vanhamme (2023).	Assessment of conceptual coherence among testing, standards, maintenance, warranty, and parts support within safety, with clear boundaries from quality and sustainability.
Sustainability	Resource, energy, disassembly, repair, related to durability, and end of life implications across the product lifecycle.	Bocken et al. (2016); Isaksson Hallstedt et al. (2023); Du & Ayob (2024); Mesa et al. (2022).	Assessment of lifecycle relevance and clear conceptual distinction from general quality attributes.
Quality	Dependable long-term performance, specifications, material fitness, reliability, durability, and conformance.	Stylidis et al. (2020); Bayram et al. (2024); Körkemeyer et al. (2025).	Assessment of performance coherence and the separation of material fitness from environmental material implications.
Commerciality	Market related cues that influence value, risk,	Beneke et al. (2013); Ismagilova et al.	Assessment of brand, price, reviews, recommendation, information, and

Criterion	Conceptual role	Literature support	Validation focus
	credibility, information sufficiency, service support, and purchase intention.	(2020); Abdul Jamil & Abu Bakar (2025).	service as market related inputs to design decision support, distinct from intrinsic technical quality.
Innovation	Value creation through benefit, differentiation, perceived newness, novelty, and symbolic market positioning.	Homburg et al. (2015); Sung et al. (2022); Wang et al. (2022); Abdul Jamil & Abu Bakar (2025).	Assessment of benefit and novelty as distinct from prestige, rarity, and exclusivity, which are also influenced by market positioning.

(Source: Authors' synthesis)

3 METHOD

3.1 Research Design and Object of Validation

A quantitative expert content validation design was used to evaluate the SPDM. The study evaluated the content architecture of the model rather than creating a new consumer ranking. The model contains 37 indicators across seven criteria. The indicator order had already been established from user evidence, so expert scores were analysed as evidence of relevance and conceptual adequacy.

The design separates validation from reconstruction. An indicator was retained when it met the predetermined I-CVI threshold, even when another indicator received a higher I-CVI. A high I-CVI did not justify moving an indicator upward in the priority order derived from user evidence. This procedure preserves the empirical meaning of the user layer while expert judgement strengthens terminology and conceptual boundaries.

Table 4 SPDM Validation Architecture

SUSTAINABILITY FOUNDATIONS	PERCEIVED PRODUCT QUALITY	PURCHASE INTENTION
Social • Environmental • Economic	How users interpret overall product excellence	Willingness to consider or choose the product
PRIORITY DERIVED FROM USER EVIDENCE	**EXPERT CONTENT VALIDATION**	**DESIGN DECISION SUPPORT**
Priority order within 7 criteria	I-CVI • S-CVI/Ave • Modified kappa • conceptual refinement	FORM • FUNCTION • SAFETY • SUSTAINABILITY • QUALITY • COMMERCIALITY • INNOVATION

(Source: Authors' synthesis)

3.2 Expert Panel and Recruitment

The validation panel comprised 30 experts from product design and industry practice. Eligibility required at least five years of relevant professional experience. This experience supported evaluation of the criteria in relation to product development, usability, sustainability, quality, market context, compliance, and lifecycle support. The panel represented more than one professional perspective, which reduced dependence on a single design specialism.

Questionnaires were administered face to face between May and July 2025. The researcher handed the questionnaire directly to each expert for in person completion. Direct administration provided a consistent setting for presenting the study instructions and allowed procedural clarification when required. Methodological research shows that face to face questionnaire administration can support

accurate question interpretation and response completeness when clarification is controlled and nonleading (Conrad & Schober, 2021; Edwards, 2010). Participation was voluntary, and informed consent was obtained before questionnaire completion. The study collected numeric relevance judgements only, with no interview responses used as research data. The analysis focuses on overall panel agreement, while subgroup comparison by discipline, sector, or years of experience lies outside the available dataset.

3.3 Validation Instrument and Rating Procedure

Each expert evaluated the relevance of every indicator using a four-point scale: 1 = not relevant, 2 = somewhat relevant / major revision required, 3 = relevant / minor revision required, and 4 = highly relevant. Scores of 3 or 4 were classified as relevant for CVI calculation, while scores of 1 or 2 were classified as lower relevance. The four-point structure removes a neutral midpoint and requires a directional relevance judgement, consistent with established CVI practice (Lynn, 1986; Polit & Beck, 2006; Yusoff, 2019).

The validation form retained the structure of seven criteria and the indicator labels so that the expert task remained focused on content representation. Experts were not instructed to create a new ranking. This distinction is necessary because content validity concerns relevance and representativeness, whereas ranking concerns relative priority.

3.4 Content Validity Analysis

For each indicator, I-CVI was calculated as A/N, where A is the number of experts assigning a rating of 3 or 4 and N = 30. Thirty relevant ratings produced an I-CVI of 1.00; 29 produced 0.97; 28 produced 0.93; 27 produced 0.90; 26 produced 0.87; 25 produced 0.83; and 24 produced 0.80. The predetermined retention threshold was I-CVI $\geq$ 0.78, consistent with widely used content validity guidance for panels larger than six experts (Polit et al., 2007; Yusoff, 2019).

S-CVI/Ave was calculated separately for each criterion by averaging the I-CVI values of its indicators. The overall mean I-CVI across all 37 indicators was also calculated to summarise the relevance profile of the model. Universal agreement becomes increasingly strict as panel size increases; therefore, S-CVI/Ave was used as the primary scale level summary rather than S-CVI/UA (Zamanzadeh et al., 2015).

Modified kappa (k*) was calculated to adjust each I-CVI for chance agreement. The probability of chance agreement was calculated as Pc = [N! / A!(N-A)!] × 0.5^N, and k* = (I-CVI - Pc) / (1 - Pc). Following established content validity interpretation, k* values above 0.74 were classified as excellent agreement (Polit et al., 2007; Zamanzadeh et al., 2015). With N = 30 and high observed agreement counts, Pc was very small and k* remained close to I-CVI. Modified kappa was therefore used as a robustness check against chance agreement, not as independent evidence of construct validity.

3.5 Analytical Strategy and Decision Rules

The analysis followed a retain, validate, and clarify logic. Threshold compliance determined retention, differences at criterion and item levels identified consensus patterns, and indicators with lower consensus received conceptual review. Table 5 summarises the analytical protocol.

Table 5 Validation Protocol and Decision Rules

Analytical step	Statistic / evidence	Decision rule	Interpretive purpose
Item relevance	I-CVI	Retain when I-CVI ≥ 0.78.	Measures the relevance of each indicator against the retention criterion.
Criterion relevance	S-CVI/Ave	Compare criterion averages; no reranking of user priorities.	Shows concentration or dispersion of professional agreement by criterion.
Chance correction	Modified kappa	k* > 0.74 interpreted as excellent agreement.	Confirms item agreement beyond chance.
Diagnostic review	Relative I-CVI pattern + conceptual literature	Review items with comparatively lower agreement.	Identifies ambiguous wording, overlapping boundaries, or context dependence.
Integration	User priority + expert relevance	Preserve user order while refining conceptual boundaries identified by experts.	Combines experiential legitimacy with professional defensibility.

(Source: Authors' analytical procedure)

4 RESULTS

4.1 Overall Content Validity

All 37 indicators met the predetermined I-CVI retention criterion. I-CVI values ranged from 0.80 to 1.00, and the overall mean I-CVI was 0.94. Criterion level S-CVI/Ave values ranged from 0.88 to 0.98. FUNCTION recorded the strongest criterion level agreement (0.98), followed by QUALITY (0.97), FORM (0.95), SAFETY (0.93), SUSTAINABILITY (0.93), COMMERCIALITY (0.91), and INNOVATION (0.88). The results support retention of the seven criteria while showing clear differences in the level of conceptual consensus.

Table 6 Criterion Level Content Validity Summary

Criterion	No. of indicators	S-CVI/Ave	Lowest I-CVI	Highest I-CVI	Items with unanimous relevance	Items retained
Form	5	0.95	0.90	1.00	1	5
Function	5	0.98	0.97	1.00	2	5
Safety	5	0.93	0.87	1.00	1	5
Sustainability	5	0.93	0.83	1.00	1	5
Quality	5	0.97	0.93	1.00	2	5
Commerciality	7	0.91	0.83	0.97	0	7
Innovation	5	0.88	0.80	0.97	0	5

(N = 30; Source: Authors' analysis)

Table 6 shows that the strongest consensus is concentrated in criteria closely connected with observable product use and performance. FUNCTION, QUALITY, and FORM all exceed 0.95. COMMERCIALITY and INNOVATION show greater dispersion because their indicators depend more strongly on market context, symbolic meaning, and interpretation. This pattern identifies the criteria that require clearer operational language rather than structural deletion.

4.2 Indicator Level Content Validity and Modified Kappa

Table 7 reports all 37 indicators, relevant rating counts, I-CVI, modified kappa, and the resulting decision. Every modified kappa value exceeded 0.79 and fell within the excellent range. The chance corrected results confirm that correction for chance does not change the item level retention pattern.

Table 7 Indicator Level Content Validity and Chance Corrected Agreement

Criterion	Code	Indicator	Relevant ratings	I-CVI	Modified kappa	Agreement	Decision
Form	D1	Shape	30/30	1.00	1.00	Excellent	Retain
Form	D2	Appearance	29/30	0.97	0.97	Excellent	Retain
Form	D3	Color	29/30	0.97	0.97	Excellent	Retain
Form	D4	Product Finishing	28/30	0.93	0.93	Excellent	Retain
Form	D5	Texture	27/30	0.90	0.90	Excellent	Retain
Function	E1	Easy to Understand	29/30	0.97	0.97	Excellent	Retain
Function	E2	Ergonomic	30/30	1.00	1.00	Excellent	Retain
Function	E3	Comfortable	29/30	0.97	0.97	Excellent	Retain
Function	E4	Convenient / User Friendly / Easy to Use	30/30	1.00	1.00	Excellent	Retain
Function	E5	Performance	29/30	0.97	0.97	Excellent	Retain
Safety	F1	Spare Part Availability	26/30	0.87	0.87	Excellent	Retain
Safety	F2	Test Results	29/30	0.97	0.97	Excellent	Retain
Safety	F3	Warranty Provisions	27/30	0.90	0.90	Excellent	Retain
Safety	F4	Maintenance	28/30	0.93	0.93	Excellent	Retain
Safety	F5	Compliance with Standards and Regulations	30/30	1.00	1.00	Excellent	Retain
Sustainability	G1	Stability	25/30	0.83	0.83	Excellent	Retain
Sustainability	G2	Material Use / Type	29/30	0.97	0.97	Excellent	Retain
Sustainability	G3	Energy Efficient	30/30	1.00	1.00	Excellent	Retain
Sustainability	G4	Assemble and Disassemble	28/30	0.93	0.93	Excellent	Retain
Sustainability	G5	Recyclability	27/30	0.90	0.90	Excellent	Retain
Quality	H1	Reliability	30/30	1.00	1.00	Excellent	Retain
Quality	H2	Durability	29/30	0.97	0.97	Excellent	Retain
Quality	H3	Product Specifications	30/30	1.00	1.00	Excellent	Retain
Quality	H4	Material Selection	28/30	0.93	0.93	Excellent	Retain
Quality	H5	Conformance	29/30	0.97	0.97	Excellent	Retain
Commerciality	I1	Brand	26/30	0.87	0.87	Excellent	Retain
Commerciality	I2	Recommendation	28/30	0.93	0.93	Excellent	Retain
Commerciality	I3	Price	25/30	0.83	0.83	Excellent	Retain
Commerciality	I4	Product Information	29/30	0.97	0.97	Excellent	Retain
Commerciality	I5	Customer Review	27/30	0.90	0.90	Excellent	Retain
Commerciality	I6	Service after Purchase	28/30	0.93	0.93	Excellent	Retain
Commerciality	I7	Online Reviews	29/30	0.97	0.97	Excellent	Retain

Criterion	Code	Indicator	Relevant ratings	I-CVI	Modified kappa	Agreement	Decision
Innovation	J1	Product Benefit	29/30	0.97	0.97	Excellent	Retain
Innovation	J2	Uniqueness	27/30	0.90	0.90	Excellent	Retain
Innovation	J3	Prestige Value	24/30	0.80	0.80	Excellent	Retain
Innovation	J4	Exclusiveness / Limited	24/30	0.80	0.80	Excellent	Retain
Innovation	J5	Novelty / Originality	28/30	0.93	0.93	Excellent	Retain

(N = 30; Source: Authors' analysis)

The distribution is concentrated at the upper end of the scale. Seven indicators achieved unanimous relevance (I-CVI = 1.00), while most remaining indicators ranged from 0.90 to 0.97. Only four indicators fell below 0.87: Stability (0.83), Price (0.83), Prestige Value (0.80), and Exclusiveness/Limited (0.80). These values met the retention criterion and were examined as diagnostic signals for closer conceptual interpretation.

4.3 Strongest Consensus Clusters

FUNCTION produced the strongest overall profile. Ergonomic and Convenient/User Friendly/Easy to Use achieved unanimous relevance, while Easy to Understand, Comfortable, and Performance each achieved 0.97. The narrow spread shows consistent expert interpretation of function as a multidimensional interaction between users and products rather than as technical performance alone. The results also support a distinction between comprehension and ease of use. Ease of operation after learning does not guarantee immediate comprehension, and immediate comprehension does not guarantee physical ease of use.

QUALITY showed a similarly coherent profile. Reliability and Product Specifications achieved unanimous relevance, Durability and Conformance achieved 0.97, and Material Selection achieved 0.93. The criterion combines product performance, consistency, specification clarity, and material fitness. The main issue is the conceptual boundary between Material Selection under QUALITY and Material Use/Type under SUSTAINABILITY, not insufficient expert agreement.

FORM also demonstrated strong consensus. Shape achieved 1.00, Appearance and Color 0.97, Product Finishing 0.93, and Texture 0.90. Texture recorded the lowest value within FORM. This result is consistent with the literature because texture is interpreted through visual, material, and tactile qualities, and its relevance depends on the mode of product interaction. Md. Nazri et al. (2025) emphasise multisensory and emotional product experience, while Alias and Che Din (2025) show that specific visual design elements influence purchase decisions. Texture therefore remains valid, with its operational definition tied to the intended sensory channel and product context.

4.4 Criterion Specific Interpretation

SAFETY: SAFETY achieved S-CVI/Ave = 0.93. Compliance with Standards and Regulations achieved unanimous relevance, followed by Test Results (0.97), Maintenance (0.93), Warranty Provisions (0.90), and Spare Part Availability (0.87). The pattern separates direct safety evidence from lifecycle support. Standards and tests provide explicit safety evidence during validation or market entry, whereas maintenance, warranty, and parts availability support continued safe performance over time. This lifecycle interpretation is consistent with durability and repair literature. Mesa et al. (2022) link repair and maintenance to product life extension, Amend et al. (2022) show that design and service support influence repair behaviour, and Munten and Vanhamme (2023) demonstrate that repairability signals durability and quality. The literature and item definition place Spare Part Availability at the boundary between safety, serviceability, and sustainability.

SUSTAINABILITY: SUSTAINABILITY achieved S-CVI/Ave = 0.93. Energy Efficient achieved 1.00, Material Use/Type 0.97, Assemble and Disassemble 0.93, Recyclability 0.90, and Stability 0.83. Energy, material, disassembly, and recyclability map directly onto established circular design and lifecycle concerns (Bocken et al., 2016; Isaksson Hallstedt et al., 2023; Mesa et al., 2022), while Du and Ayob (2024) reinforce sustainability as part of product design meaning. The term Stability covers physical stability, dimensional stability, system stability, material stability, and long-term performance consistency. Within SUSTAINABILITY, the operational definition focuses on service life continuity and the avoidance of premature failure.

COMMERCIALITY: COMMERCIALITY achieved S-CVI/Ave = 0.91. Product Information and Online Reviews achieved 0.97, Recommendation and Service after Purchase 0.93, Customer Review 0.90, Brand 0.87, and Price 0.83. The wider spread is consistent with the contextual nature of market related evaluation. Meta analytic evidence shows that electronic word of mouth and the usefulness and credibility of digital review information affect intention to buy (Ismagilova et al., 2020), while Abdul Jamil and Abu Bakar (2025) show that design development interacts with public perception, brand image, and purchase intention. These indicators represent the market interface of the model rather than intrinsic product properties. Price is relational because its meaning depends on segment, alternatives, purchasing power, and the value communicated by the complete product proposition.

INNOVATION: INNOVATION achieved the lowest S-CVI/Ave (0.88), although all five indicators met the retention criterion. Product Benefit achieved 0.97, Novelty/Originality 0.93, Uniqueness 0.90, while Prestige Value and Exclusiveness/Limited each achieved 0.80. The pattern separates innovation as meaningful value creation from innovation as symbolic positioning. Sung et al. (2022) show that objective design newness and consumers' subjective perception of newness are distinct, while Wang et al. (2022) demonstrate that exclusivity and rarity operate as separate market constructs. Abdul Jamil and Abu Bakar (2025) similarly connect innovation-oriented design development with acceptance rather than treating novelty alone as sufficient. Prestige and exclusivity also result from branding, scarcity, and status positioning independent of technical or functional novelty.

4.5 Diagnostic Matrix for Indicators with Lower Consensus

All four indicators with lower expert agreement remained above the retention threshold. The analysis therefore used these results to identify the conceptual source of disagreement rather than treating them as failed items. Table 8 presents the corresponding refinement actions.

Table 8 Diagnostic Interpretation of Indicators with Lower Consensus

Indicator	I-CVI	Diagnostic issue	Refinement action	Status
Stability	0.83	The term covers physical stability, dimensional stability, system stability, and long term performance consistency.	Define specifically as sustained safe and intended functional performance across expected use conditions and service life when used under Sustainability.	Retain and clarify
Price	0.83	Variable that depends on market segment, positioning, promotion, and purchasing power.	Define as perceived price appropriateness/value within the target market rather than as a fixed product attribute.	Retain and contextualise
Prestige Value	0.80	Can reflect status consumption or brand symbolism rather than innovation itself.	Frame as symbolic value created by perceived advancement or differentiation; avoid treating prestige as evidence of technical innovation.	Retain with theoretical boundary
Exclusiveness / Limited	0.80	Scarcity also results from marketing strategy rather than novelty.	Frame as perceived differentiation or scarcity-based market positioning and distinguish it from originality.	Retain with theoretical boundary

(Source: Authors' analysis)

4.6 Refined Operational Definitions for Conceptually Adjacent Indicators

The validation pattern also exposed boundaries between indicators that are individually valid but conceptually adjacent. The most important example is Material Selection under QUALITY versus Material Use/Type under SUSTAINABILITY. Table 8 formalises definitions that preserve their distinct analytical roles.

Table 9 Refined Operational Definitions for Selected Indicators

Indicator	Criterion	Refined operational definition	Boundary from adjacent concepts
Stability	Sustainability	Capacity of the product to maintain safe, consistent, and intended functional performance across expected use conditions and service life.	Extends beyond resistance to tipping by defining long term functional continuity throughout the intended service life.
Material Use / Type	Sustainability	Environmental and resource implications of material type and quantity, including renewability, recycled content, toxicity, circularity, and resource intensity.	Distinct from fitness for purpose or finish quality.
Material Selection	Quality	Suitability of the selected material for required performance, durability, finish, reliability, and conformance.	Distinct from the environmental burden of the material choice.
Commerciality	Commerciality	Market related conditions that shape perceived value, credibility, risk, information sufficiency, service support, and purchase desirability.	Represents the market interface of the model and remains distinct from intrinsic technical product properties.
Prestige Value	Innovation	Symbolic value associated with perceived advancement, differentiation, or status generated by the product proposition.	Represents symbolic value linked to perceived advancement or differentiation and remains separate from evidence of technical novelty.
Exclusiveness / Limited	Innovation	Perceived scarcity or differentiated availability that contributes to market distinctiveness.	Separate from novelty/originality and from brand prestige.

(Source: Authors' synthesis from validation results)

5 DISCUSSION

5.1 Strength of the Seven Criteria Architecture

The central finding is that all 37 indicators exceeded the retention threshold and every criterion achieved an S-CVI/Ave of at least 0.88, while the level of expert agreement differed across items and criteria. This combination supports the architecture and preserves useful diagnostic variation. Threshold compliance does not represent complete validation. Almanasreh et al. (2019) emphasise that content validity depends on domain representation and conceptual specification as well as numerical agreement. In this study, the variation from 0.80 to 1.00 identifies where professional interpretation depends more strongly on context and where definitions require greater precision.

FUNCTION, QUALITY, and FORM recorded the strongest results because these criteria are directly connected with observable product interaction, performance, and sensory cues. Bayram et al. (2024) identify functionality, technical attributes, longevity, design, and customer value as recurring

dimensions of product quality in circular economy research. Mesa et al. (2022) further show that durability and continued functional performance are central to product life extension, while Md. Nazri et al. (2025) and Alias and Che Din (2025) show how sensory and visual design features shape product experience and purchase evaluation. The convergence between the validation results and these studies supports the high expert agreement and confirms that sustainable value must operate together with usability, dependable performance, and perceptual quality.

COMMERCIALITY and INNOVATION recorded wider dispersion because their indicators depend more strongly on context and interpretation. Brand, price, reviews, prestige, exclusivity, and novelty vary with product category, market segment, social meaning, and competitive reference. Ismagilova et al. (2020) show that the effect of digital review information on purchase intention depends on message usefulness, credibility, and related eWOM mechanisms; Sung et al. (2022) show that perceived design newness differs from objective novelty; and Wang et al. (2022) distinguish exclusivity from rarity. The lower S-CVI/Ave values for these market related criteria are therefore consistent with evidence that their meanings depend on context. Their role in the SPDM requires explicit operational boundaries.

5.2 User Priority and Expert Relevance as Complementary Evidence

A major methodological contribution of this study is the separation of priority derived from user evidence from relevance established through expert evidence. The two forms of evidence answer different questions and remain distinct throughout model use. User priority indicates relative importance from the perspective of intended users, while expert content validity examines the conceptual fit of an indicator within its proposed criterion. This distinction is consistent with content validity methodology, which concerns relevance and representativeness rather than preference ranking (Almanasreh et al., 2019; Polit & Beck, 2006). It also aligns with consumer evidence showing that design priorities and purchase related judgements are empirically meaningful (Mohd Rashid et al., 2026). The expert layer therefore evaluates the conceptual structure while preserving the user layer.

The sustainability criterion shows why this separation matters. Energy Efficient achieved unanimous expert relevance, while Stability achieved 0.83. Using I-CVI to reorder the criterion converts a validity statistic into a consumer priority statistic and alters the empirical meaning of the framework. The appropriate procedure is to retain the order derived from user evidence and use expert evidence to clarify the meaning of Stability. The same logic applies to Price, Prestige Value, and Exclusiveness/Limited. Expert disagreement is therefore used as evidence for construct diagnosis rather than as an automatic deletion rule.

This dual evidence logic also provides a transferable approach for other user centred frameworks. User evidence provides direct evidence of intended user priorities. Expert evidence establishes professional support for conceptual relevance, terminology, feasibility, and boundary coherence. Convergence strengthens the evidence base, while divergence identifies where theory, wording, or context requires refinement.

Table 10 Integration of User and Expert Evidence in the Validated SPDM

Evidence layer	Analytical focus	Evidence provided	Scope boundary	Role in final model
User priority	Relative importance to intended users.	Relative priority within each criterion.	Conceptual placement and professional defensibility require expert evidence.	Preserved ordering / prioritisation.
Expert relevance	Conceptual representation of the proposed criterion.	Content relevance and representativeness.	Consumer priority, causality, predictive performance, and actual purchase behaviour require separate evidence.	Retention and criterion validation.
Modified kappa	Agreement after correction for chance.	Strength of item agreement beyond chance.	Practical importance and market effect size require separate analysis.	Robustness check.
Conceptual refinement	Areas with greater interpretive variation.	Need for clearer definitions and boundaries.	Deletion or reranking requires evidence beyond comparative expert variation.	Operational clarification.
Integrated framework	Joint use of user and expert evidence in design.	A validated design reference combining user legitimacy with professional defensibility.	Application across sectors and cultures requires further validation.	Decision support during the early design stage.

(Source: Authors' synthesis)

5.3 Conceptual Contributions to Sustainable Product Design

The findings contribute to sustainable product design in four conceptual ways. First, they support sustainability as a design criterion embedded alongside form, function, safety, quality, commerciality, and innovation rather than as a separate compliance layer. This is consistent with circular quality literature showing relationships among functionality, longevity, customer value, design, and environmental aspects (Bayram et al., 2024) and with studies that integrate sustainability with product meaning and human centred design (Du & Ayob, 2024; Mohd Hasbullah et al., 2025). Second, the results make boundary management an explicit part of model validation. Material Selection and Material Use/Type remain distinct when performance fitness is separated from environmental resource implications. Third, the study clarifies the market interface of sustainable design by retaining commerciality as a distinct market related criterion. Fourth, lower consensus is used as analytical evidence to identify constructs that depend more strongly on context, symbolism, or terminology.

The material boundary is important because circular product quality and conventional performance quality overlap in practice. Mesa et al. (2022) connect material and durability decisions with product life extension, while Bayram et al. (2024) show that circular economy quality literature includes both technical and environmental aspects. The SPDM therefore treats Material Selection as fitness for required performance and Material Use/Type as the environmental and resource consequence of the material decision. This distinction prevents double counting while recognising that one material decision influences several design objectives.

The market boundary is equally important. COMMERCIALITY does not redefine price, reviews, brand, or recommendation as technical properties. It recognises that product value must be communicated, trusted, and supported before lifecycle benefits are realised in use. Meta analytic eWOM evidence confirms that digital review information affects intention to buy (Ismagilova et al., 2020), while Abdul Jamil and Abu Bakar (2025) link design development with public perception and purchase intention. Sustainable design models therefore require a market interface that remains analytically distinct from function, quality, and innovation.

5.4 Implications for Product Designers and Industry Practitioners

For product designers, the validated SPDM functions as a structured review framework during requirement definition, concept evaluation, and design development. FORM guides visible and tactile quality cues; FUNCTION guides comprehension, ergonomics, comfort, usability, and performance; SAFETY links compliance evidence with lifecycle support; SUSTAINABILITY focuses on energy, materials, disassembly, service life, and end of life concerns; QUALITY focuses on reliable and conforming performance; COMMERCIALITY captures market related value and support; and INNOVATION captures benefit, differentiation, novelty, and symbolic value.

For industry practitioners, the framework supports multidisciplinary discussion. Engineering, industrial design, marketing, sustainability, quality assurance, regulatory, procurement, and after sales teams often use different definitions of product success. A common framework makes trade-offs explicit. Trade-offs arise when a material decision improves environmental performance but weakens finish quality, when an innovative feature increases novelty but reduces ease of use, or when a lower cost design improves price competitiveness but reduces durability or repairability. The SPDM does not mechanically resolve these trade-offs. It provides an evidence-based vocabulary for identifying them during early design decisions.

The model is best used as a decision support architecture rather than as a fixed scoring checklist. Teams retain priorities derived from user evidence while using the validated criteria to structure design reviews, document trade-offs, and identify evidence gaps. In this role, the framework supports design reasoning without imposing one universal weighting scheme across all product categories or markets.

5.5 Critical Robustness and Boundary Conditions

Three methodological points define the scope of the strong content validity profile. First, the four point expert ratings were converted into relevant (3 or 4) and lower relevance (1 or 2) categories for I-CVI. This procedure improves interpretive clarity but removes the distinction between 'relevant with minor revision' and 'highly relevant'. The same I-CVI therefore does not guarantee the same endorsement intensity. Future validation should report the complete rating distribution alongside CVI so that relevance strength and disagreement remain visible.

Second, modified kappa was interpreted within its statistical role. With 30 experts and between 24 and 30 relevant ratings, the probability of chance agreement is extremely small; accordingly, k* is numerically close to I-CVI. Its role in this study is to confirm that chance correction does not change the retention decisions. It is not a second independent validity test. Construct validity, discriminant validity, predictive validity, and practical effectiveness require separate evidence.

Third, S-CVI/Ave averages item scores, so a strong criterion average does not rule out a comparatively ambiguous indicator. The present analysis addresses this issue by retaining item level diagnostics and by not treating criterion averages as proof of homogeneity. This is particularly important for INNOVATION and COMMERCIALITY, where the literature shows that newness, exclusivity, price, and digital review information depend on context (Ismagilova et al., 2020; Sung et al., 2022; Wang et al., 2022). The strength of the SPDM therefore depends on clear boundary definitions rather than a claim that every criterion remains equally stable across all contexts.

5.6 Comparison with Previous Research

Table 11 compares the main validation findings with prior empirical and review evidence. The strongest SPDM criteria correspond with established dimensions of product experience and quality, while the criteria with lower consensus correspond with constructs that external research describes as dependent on market context, symbolism, or perception. This convergence supports the interpretation

that lower expert agreement reflects conditional construct meaning rather than random variation.

Table 11 Comparison of Main Findings with Previous Research

Current finding	Related literature	Interpretation	Design implication
FUNCTION achieved the highest S-CVI/Ave (0.98).	Golder et al. (2012); Bayram et al. (2024); Mohd Rashid et al. (2026).	Functionality is central to quality experience, but contemporary evidence treats function as usability, comprehension, ergonomics, and performance rather than performance alone.	Keep functional criteria user focused and measurable; do not reduce function to technical output.
QUALITY achieved S-CVI/Ave = 0.97.	Stylidis et al. (2020); Bayram et al. (2024); Mesa et al. (2022); Körkemeyer et al. (2025).	Reliability, durability, specifications, and conformance support long term value and circularity, but material quality and environmental material impact remain distinct questions.	Integrate quality early with sustainability while maintaining construct boundaries.
SAFETY achieved S-CVI/Ave = 0.93, with Spare Part Availability lower at 0.87.	Mesa et al. (2022); Amend et al. (2022); Munten & Vanhamme (2023).	Lifecycle support contributes to safe and durable use, but spare parts and maintenance also belong to repairability and serviceability.	Define safety as continued safe operation and make connections across criteria links explicit.
Material Selection and Material Use/Type require boundary clarification.	Bayram et al. (2024); Isaksson Hallstedt et al. (2023); Mesa et al. (2022).	Performance fitness and environmental resource implications are related but analytically distinct.	Separate material suitability from resource, toxicity, circularity, and environmental consequences.
COMMERCIALITY remains valid but dependent on context.	Beneke et al. (2013); Ismagilova et al. (2020); Abdul Jamil & Abu Bakar (2025).	Market cues shape perceived value, risk, credibility, and willingness to buy, but their influence varies by channel, segment, and information context.	Use commerciality as a market interface criterion rather than as intrinsic product quality.
INNOVATION shows the widest conceptual variation (S-CVI/Ave = 0.88).	Homburg et al. (2015); Sung et al. (2022); Wang et al. (2022); Abdul Jamil & Abu Bakar (2025).	Meaningful benefit and perceived newness are closer to innovation than prestige or scarcity, which can arise from positioning.	Define symbolic innovation dimensions carefully and do not equate exclusivity with technical novelty.
User priority and expert relevance are complementary, not interchangeable.	Polit & Beck (2006); Almanasreh et al. (2019); Mohd Rashid et al. (2026).	Different evidence sources answer different validation questions; content relevance cannot substitute for user priority.	Preserve user ordering and use expert evidence for relevance, robustness, and clarification.

(Source: Authors' synthesis)

Table 12 Critical Evidence Triangulation Across the Seven SPDM Criteria

Criterion	Validation signal	External evidence	Critical qualification	Implication for SPDM use
FORM	S-CVI/Ave = 0.95	Md. Nazri et al. (2025); Alias & Che Din (2025)	Visual and tactile cues matter, but their salience depends on the product and sensory context.	Specify the sensory channel and context for texture, colour, and finishing.
FUNCTION	S-CVI/Ave = 0.98	Bayram et al. (2024); Mohd Rashid et al. (2026)	Very high agreement demonstrates criterion coherence; causal effects on purchase and sustainability require separate empirical testing.	Use as a core review domain; validate predictive relationships separately.
SAFETY	S-CVI/Ave = 0.93	Mesa et al. (2022); Amend et al. (2022); Munten & Vanhamme (2023)	Maintenance and spare parts cross into repairability and lifecycle support.	Retain lifecycle support but make overlap with sustainability explicit.
SUSTAINABILITY	S-CVI/Ave = 0.93	Bocken et al. (2016); Isaksson Hallstedt et al. (2023); Du & Ayob (2024)	Stability is valid but semantically broad and risks duplication with quality or safety.	Define stability as service life continuity under expected use conditions.
QUALITY	S-CVI/Ave = 0.97	Stylidis et al. (2020); Bayram et al. (2024); Mesa et al. (2022)	Material fitness and material environmental burden are related but distinct.	Keep Material Selection separate from Material Use/Type.
COMMERCIALITY	S-CVI/Ave = 0.91	Ismagilova et al. (2020); Abdul Jamil & Abu Bakar (2025)	Price, reviews, and brand are relational and dependent on the information channel.	Treat as a market interface and contextualise by target segment.
INNOVATION	S-CVI/Ave = 0.88	Sung et al. (2022); Wang et al. (2022)	Prestige and exclusivity represent symbolic positioning and remain analytically distinct from technical or functional innovation.	Distinguish meaningful newness from scarcity/status positioning.

(Source: Authors' synthesis)

5.7 Limitations and Future Research

The findings are interpreted within the scope of content validation. I-CVI, S-CVI/Ave, and modified kappa establish expert relevance and agreement; they do not establish construct validity, discriminant validity, criterion validity, predictive validity, causality, or the effectiveness of the SPDM in live design projects. Converting the four response categories into relevant and lower relevance groups also removes information about endorsement intensity. In addition, the expert panel is analysed as one group because the available dataset does not support reliable stratification by professional discipline, sector, or experience level. The present evidence is therefore strong for content representation and clearly defines the next stages required for broader validation.

The model was developed within a specific product design and consumer context. The seven criteria support wider application, but generalisation to medical devices, mobility systems, furniture, integrated digital and physical products, industrial equipment, or other cultural markets requires further testing. The meaning of some indicators, especially commerciality and innovation, varies across product categories.

Future research should test the SPDM with independent expert panels, additional product categories, and different cultural settings; preserve the full four category rating distribution; and examine convergent and discriminant relationships among the criteria. Predictive studies should examine the extent to which evaluations based on the SPDM explain concept selection, user acceptance, perceived quality, or purchase intention beyond simpler design checklists. Future field studies should compare teams using the SPDM with teams using conventional design development procedures. Because the current data come from questionnaire responses, future qualitative studies should investigate the reasons for lower consensus as a separate empirical phase instead of deriving those reasons from the present numeric data.

6 CONCLUSION

This study provides content validity evidence for a Sustainable Product Design Model containing 37 indicators and evaluated by 30 experienced product designers and industry practitioners. Every indicator exceeded the predetermined I-CVI threshold of 0.78, with values ranging from 0.80 to 1.00 and an overall mean of 0.94. Criterion level S-CVI/Ave values ranged from 0.88 to 0.98, and modified kappa indicated excellent agreement after correction for chance across all indicators.

The results support retention of the seven criteria while showing meaningful differences in conceptual consensus. FUNCTION, QUALITY, and FORM were especially coherent and align with literature that places functionality, durability, sensory cues, and perceived quality at the centre of product experience. COMMERCIALITY and INNOVATION showed greater variation, consistent with evidence that digital review information, newness, exclusivity, and symbolic value depend on context. Stability, Price, Prestige Value, and Exclusiveness/Limited therefore remain valid and require clear operational boundaries.

The main contribution is both methodological and substantive. Priority derived from user evidence and relevance established through expert evidence are treated as complementary forms of evidence rather than competing rankings, and content validity statistics are interpreted within clearly stated limits. Preserving user priority maintains alignment with intended users. Expert validation establishes professional support and identifies definitions that require refinement. External literature provides a third interpretive layer for assessing the consistency between the observed consensus pattern and established theory. The validated SPDM therefore provides a stronger basis for structured sustainable product design decisions in the early design stage and defines clear directions for subsequent construct, predictive, and field validation.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the product designers and industry practitioners who contributed their professional judgements to the expert validation stage of this research. The authors also acknowledge the Faculty of Arts and Design, Universiti Teknologi MARA, for supporting the academic development of the study.

FUNDING

This research was funded by the authors.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptual development, review, analysis, and preparation of this paper.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

DATA AVAILABILITY

The aggregated validation data supporting the findings of this study are reported in the tables within this article. Individual expert responses are not reproduced in order to preserve participant confidentiality.

REFERENCES

Aaker, D. A., & Joachimsthaler, E. (2000). Brand leadership: Building assets in an information economy. Free Press.

Abdul Jamil, A. R., & Abu Bakar, M. F. (2025). Design development for product extension in the context of public acceptance in Malaysia. Idealogy Journal, 10(1), 206-217. https://doi.org/10.24191/idealogy.v10i1.767

Agarwal, S., & Teas, R. K. (2004). Cross-national applicability of a perceived risk-value model. Journal of Product & Brand Management, 13(4), 242-256. https://doi.org/10.1108/10610420410546952

Alias, S. S., & Che Din, S. (2025). Analysis of design elements on soda drink packaging in influencing youth consumer purchase decision. Idealogy Journal, 10(2), 402-412. https://doi.org/10.24191/idealogy.v10i2.793

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

Amend, C., Revellio, F., Tenner, I., & Schaltegger, S. (2022). The potential of modular product design on repair behavior and user experience: Evidence from the smartphone industry. Journal of Cleaner Production, 367, 132770. https://doi.org/10.1016/j.jclepro.2022.132770

Bayram, B., Deserno, L., & Greiff, K. (2024). Product quality in the circular economy: A systematic review of its definition and contexts in scientific literature. Circular Economy and Sustainability, 4, 2713-2747. https://doi.org/10.1007/s43615-024-00396-0

Beneke, J., Flynn, R., Greig, T., & Mukaiwa, M. (2013). The influence of perceived product quality, relative price and risk on customer value and willingness to buy: A study of private label merchandise. Journal of Product & Brand Management, 22(3), 218-228. https://doi.org/10.1108/JPBM-02-2013-0262

Bocken, N. M. P., de Pauw, I., Bakker, C., & van der Grinten, B. (2016). Product design and business model strategies for a circular economy. Journal of Industrial and Production Engineering, 33(5), 308-320. https://doi.org/10.1080/21681015.2016.1172124

Conrad, F. G., & Schober, M. F. (2021). Clarifying question meaning in standardized interviews can improve data quality even though wording may change: A review of the evidence. International Journal of Social Research Methodology, 24(2), 203-226. https://doi.org/10.1080/13645579.2020.1824627

Creusen, M. E. H., & Schoormans, J. P. L. (2005). The different roles of product appearance in consumer choice. Journal of Product Innovation Management, 22(1), 63-81. https://doi.org/10.1111/j.0737-6782.2005.00103.x

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

Du, S., & Ayob, S. (2024). From aesthetics to design: Sustainability of wabi-sabi in ceramic product design in the contemporary context. Idealogy Journal, 9(1), 7-15. https://doi.org/10.24191/idealogy.v9i1.530

Edwards, P. (2010). Questionnaires in clinical trials: Guidelines for optimal design and administration. Trials, 11, Article 2. https://doi.org/10.1186/1745-6215-11-2

Golder, P. N., Mitra, D., & Moorman, C. (2012). What is quality? An integrative framework of

processes and states. Journal of Marketing, 76(4), 1-23. https://doi.org/10.1509/jm.09.0416
Homburg, C., Schwemmle, M., & Kuehnl, C. (2015). New product design: Concept, measurement, and consequences. Journal of Marketing, 79(3), 41-56. https://doi.org/10.1509/jm.14.0199
Isaksson Hallstedt, S., Villamil, C., Lövdahl, J., & Wallin Nylander, J. (2023). Sustainability Fingerprint - guiding companies in anticipating the sustainability direction in early design. Sustainable Production and Consumption, 37, 424-442. https://doi.org/10.1016/j.spc.2023.03.015
Ismagilova, E., Slade, E. L., Rana, N. P., & Dwivedi, Y. K. (2020). The effect of electronic word of mouth communications on intention to buy: A meta-analysis. Information Systems Frontiers, 22, 1203-1226. https://doi.org/10.1007/s10796-019-09924-y
Körkemeyer, J. A., Brings, H., Montavon, B., & Schmitt, R. H. (2025). Multidimensional perceived quality: Extended level model and case study including sustainability as a quality dimension in the perception of plastic packaging. CIRP Annals, 74(1), 645-650. https://doi.org/10.1016/j.cirp.2025.03.017
Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382-385. https://doi.org/10.1097/00006199-198611000-00017
Marcon, A., Ribeiro, J. L. D., Dangelico, R. M., de Medeiros, J. F., & Marcon, É. (2022). Exploring green product attributes and their effect on consumer behaviour: A systematic review. Sustainable Production and Consumption, 32, 76-91. https://doi.org/10.1016/j.spc.2022.04.012
Md. Nazri, H., Taharuddin, N. S., Ibrahim, N. I., & Ibrahim, N. A. (2025). Reka bentuk produk emosi: Membentuk pengalaman melalui sentuhan emosi dalam reka bentuk [Emotional product design: Shaping experiences through emotional touch in design]. Idealogy Journal, 10(1), 89-100. https://doi.org/10.24191/idealogy.v10i1.713
Mesa, J. A., Gonzalez-Quiroga, A., Fernandes Aguiar, M., & Jugend, D. (2022). Linking product design and durability: A review and research agenda. Heliyon, 8(9), e10734. https://doi.org/10.1016/j.heliyon.2022.e10734
Mohd Hasbullah, M. S., Mappon, A., Roslan, N. F., & Wan Bakar, W. N. (2025). Walking the line between HCD and SCBT: Marine waste-art. Idealogy Journal, 10(2), 19-33. https://doi.org/10.24191/idealogy.v10i2.804
Mohd Rashid, M. H. H., Shaharudin, H., & Muhammad Firdaus, N. S. A. (2026). Perceived product quality and consumers' intent to purchase sustainable products. Idealogy Journal, 11(2), 397-415. https://doi.org/10.24191/idealogy.v11i2.1024
Morwitz, V. (2014). Consumers' purchase intentions and their behavior. Foundations and Trends in Marketing, 7(3), 181-230. https://doi.org/10.1561/1700000036
Munten, P., & Vanhamme, J. (2023). To reduce waste, have it repaired! The quality signaling effect of product repairability. Journal of Business Research, 156, 113457. https://doi.org/10.1016/j.jbusres.2022.113457
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. https://doi.org/10.1002/nur.20147
Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459-467. https://doi.org/10.1002/nur.20199
Shaharudin, H. (2013). A cross-cultural study of studio-based design education with particular reference to specific Australian and Malaysian contexts [Doctoral dissertation, University of Canberra].
Sierra-Fontalvo, L., Berni, A., & Borgianni, Y. (2026). Sustainability cues in users' perception and consumption: A product design-oriented systematic literature review. Design Science, 12, e1. https://doi.org/10.1017/dsj.2025.10046
Stylidis, K., Wickman, C., & Söderberg, R. (2020). Perceived quality of products: A framework and attributes ranking method. Journal of Engineering Design, 31(1), 37-67. https://doi.org/10.1080/09544828.2019.1669769
Sung, B., Hartley, N., Vanman, E. J., & Hooshmand, R. (2022). Objective vs subjective design newness. Asia Pacific Journal of Marketing and Logistics, 34(7), 1482-1502. https://doi.org/10.1108/APJML-12-2020-0897
Tsiotsou, R. (2006). The role of perceived product quality and overall satisfaction on purchase intentions. International Journal of Consumer Studies, 30(2), 207-217.

https://doi.org/10.1111/j.1470-6431.2005.00477.x
Veryzer, R. W., & Borja de Mozota, B. (2005). The impact of user-oriented design on new product development: An examination of fundamental relationships. Journal of Product Innovation Management, 22(2), 128-143. https://doi.org/10.1111/j.0737-6782.2005.00110.x
Visser, M., Schoormans, J., & Vogtländer, J. (2018). Consumer buying behaviour of sustainable vacuum cleaners: Consequences for design and marketing. Journal of Cleaner Production, 195, 664-673. https://doi.org/10.1016/j.jclepro.2018.05.114
Wang, X., Sung, B., & Phau, I. (2022). Examining the influences of perceived exclusivity and perceived rarity on consumers' perception of luxury. Journal of Fashion Marketing and Management: An International Journal, 26(2), 365-382. https://doi.org/10.1108/JFMM-12-2020-0254
Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. Education in Medicine Journal, 11(2), 49-54. https://doi.org/10.21315/eimj2019.11.2.6
Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. Journal of Caring Sciences, 4(2), 165-178. https://doi.org/10.15171/jcs.2015.017