

Operationalizing Istiqamah, Muraqabah, and Adab in Digital Education: Content Validation and Preliminary Psychometric Evidence

Noorhuda Abdul Karim^{1*}, Ahmad Sunawari Long², Ahmad Munawar Ismail², Nuzha Mohamed Taha¹, Siti Farhah A Aziz³

¹Faculty of Education and Humanities, UNITAR International University, Malaysia

²Research Centre for Theology and Philosophy, Faculty of Islamic Studies, Universiti Kebangsaan Malaysia, Malaysia

³School of Education and Humanities, Universiti Tun Abdul Razak (UNIRAZAK), Malaysia

HOW TO CITE:

Noorhuda Abdul Karim, Ahmad Sunawari Long, Ahmad Munawar Ismail, Nuzha Mohamed Taha, Siti Farhah A Aziz (2026). Operationalizing Istiqamah, Muraqabah, and Adab in Digital Education: Content Validation and Preliminary Psychometric Evidence. International Journal of Special Education, 41(24s), 826-839

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT:

Islamic psychospiritual constructs such as *Istiqamah*, *Muraqabah*, and *Adab* are frequently invoked in studies of Muslim learners, yet they remain under-operationalized in digital educational research. This study develops and content-validates an instrument that integrates these constructs with holistic digital well-being and holistic digital literacy for digital learning. Its distinctive contribution lies in operationalizing *Istiqamah* and *Muraqabah* as digital self-regulation and *Adab* as ethical digital comportment. An initial pool of 91 items was generated from theory and refined through a two-round expert-panel process into a 38-item instrument organized by an Input–Process–Output structure. Four experts evaluated the instrument across two rounds; the refined framework and all 38 items obtained an item-level Content Validity Index of 1.00, with scale-level indices of 1.00. Preliminary field evidence was subsequently obtained from an interim dataset of 202 eligible responses (195 complete cases). Internal consistency was satisfactory to strong for holistic digital well-being ($\alpha = .925$), digital self-regulation ($\alpha = .865$), digital *Adab* ($\alpha = .790$), and holistic digital literacy ($\alpha = .925$). The correlation matrix demonstrated excellent sampling adequacy ($KMO = .930$), with a significant Bartlett's test of sphericity, $\chi^2(703) = 5834.78$, $p < .001$. These findings provide content-valid and preliminary psychometric support for further full-sample construct validation.

Keywords: Adab, digital education, Istiqamah, Muraqabah, psychospiritual measurement.

1. Introduction

Psychological and educational constructs can only be meaningfully investigated when abstract concepts are translated into observable

indicators that adequately represent them. In case when a construct has been well-defined and its indicators are also common ones, then the requirement is relatively easier to fulfil. On the

other hand, in cases where the construct has evolved within a philosophical tradition in which categories have never been systematically measured before, the difficulty of the issue is increased. With regard to that, many existing studies dealing with Islamic psychospiritual constructs can be regarded as conceptually immature. Examples of constructs like *Istiqamah* (consistent adherence to a virtuous way), *Muraqabah* (enhanced self-consciousness due to consciousness of God's watchful eyes), and *Adab* (ethical behavior) are usually used in studies involving Muslim learners. Nevertheless, these constructs are often introduced as ideal norms but not as empirical variables. Therefore, the influence of these constructs on the results of learning processes is often taken for granted (Zafar & Malik, 2025; Nurhakim & Rahman, 2025).

However, this gap between theory and methodology becomes all the more significant amid the growing context of digital learning environments. Present research literature suggests that the experience, well-being, and behavioral involvement of learners in a digital environment cannot be fully understood without taking into account their spiritual and moral aspects, especially as online instructional modes place greater autonomous demands on both educators and students (Nair & Nesamany, 2021). In this regard, concepts of Islamic psychospirituality provide culturally relevant theoretical frameworks for understanding learner development. However, theoretical credibility is insufficient evidence to ensure empirical validity. Any assertions regarding how *Istiqamah* and *Muraqabah* function as mechanisms of digital self-regulation, and *Adab* helps in reducing digital incivility, need to be empirically confirmed. For such an assertion, systematic empirical investigation is needed. This requires measurement instruments that are both valid and theoretically defensible. Content validity, which is defined as "the degree to which test items represent the construct being measured" (Lynn, 1986; Polit & Beck, 2006), becomes one of the primary requirements for instrument construction.

The process of operationalization of psychospiritual constructs necessitates an explicit delineation of conceptual boundaries, dimensions, and indicators through which the existence of the construct can be identified. The difficulty lies in the fact that the psychospiritual constructs are inherently grounded in internal states which include intents, spirituality, and relational awareness and thus are not observable directly. Overemphasizing behavioral indicators may result in overlooking the spiritual meaning that brings importance into the concept whereas excessive abstraction of concepts may hinder any measurement efforts at all. Thus, instrument development involves striking the balance between conceptual and empirical concerns. The recently conducted research, involving the scale development of *Istiqamah* and the comparison of the concept of *Muraqabah* to the concept of mindfulness, shows that these problems can be overcome through proper conceptualization and psychometric analysis (Zafar & Malik, 2025; Nurhakim & Rahman, 2025). The above literature also emphasizes the need for developing a solid theoretical foundation before any measurements are made. This paper seeks to develop an instrument to operationalize the concepts of *Istiqamah*, *Muraqabah*, and *Adab* in the digital learning environments.

Recent research in Islamic higher education has begun to integrate Islamic ethical principles into digitally oriented competence measures, including instruments addressing responsible artificial intelligence literacy. The Islamic AI Literacy Scale developed for Indonesian Islamic higher education, for example, operationalizes awareness, usage, evaluation, and ethics through Islamic principles such as *Adab*, *amanah*, 'adl, *maslahah*, and *tabayyun* (Syamsudduha, Suwandi, & Tabrani ZA, 2026). Such developments demonstrate an increasing attention to culturally and religiously grounded educational measurement. These instruments, however, address principally the learner's competence in relation to digital or AI use. Less attention has been given to the psychospiritual mechanisms through which learners regulate themselves and sustain ethically oriented conduct

in digital learning. The present study addresses this gap by operationalizing *Istiqamah*, *Muraqabah*, and *Adab* as dimensions of digital self-regulation and ethical comportment, and by situating them within a wider model that also encompasses well-being and literacy.

The instrument is organized around three constructs corresponding to the input, process, and output of a holistic model of digital learning. The input construct, holistic digital well-being, is conceived as a multidimensional state of soundness across the physical, emotional, spiritual, intellectual, and social dimensions of the learner. Its operationalization draws on the eudaimonic tradition of psychological well-being (Ryff & Keyes, 1995) and on the balanced-development schema of the National Education Philosophy, while its spiritual dimension is rendered through the servant's relationship with the Creator as a source of inner tranquillity—a dimension for which validated Islamic well-being instruments provide precedent (Bagis, Adawiyah, Purnomo, & Sudjadi, 2024; Md Razali & Mohd Fithri, 2024).

The output construct, holistic digital literacy, is conceived as the capacity to use digital technology competently, critically, safely, and ethically. Its operationalization is anchored in the multidimensional model of digital literacy, which distinguishes operational, cognitive, and social-emotional dimensions (Ng, 2012), and in the five competence areas of the European Digital Competence Framework for Citizens (Vuorikari, Kluzer, & Punie, 2022), with technology-acceptance considerations informing, but not exhausting, the construct (Venkatesh, Thong, & Xu, 2012).

The process construct, digital character development, is the locus of the instrument's distinctive contribution. It comprises two dimensions. The first, digital self-regulation, operationalizes the social-cognitive account of self-regulated learning (Zimmerman, 2000) through the two Islamic constructs that give it motivational depth: *Istiqamah*, rendered as the disciplined constancy through which the learner sustains focus and commitment, and *Muraqabah*,

rendered as the vigilant self-monitoring through which the learner governs conduct in the awareness of being observed. While recent empirical work underscores the critical influence of self-regulated strategies on learner motivation and performance (Rifin et al., 2025), less attention has been given to the psychospiritual anchors that sustain these regulatory behaviors in online environments. The second dimension, digital *Adab*, operationalizes ethical comportment as the practice of courteous and responsible interaction in digital space, drawing on scholarship in Islamic digital ethics and on the netiquette literature (Chaudhary, 2020; Al-Khatib, 2023). Together these dimensions render measurable the character that a broader research programme positions as the mechanism connecting well-being to competence, and their operationalization is essential if that mechanism is to be empirically tested rather than merely asserted (Shahwar & Dhar, 2024; Bulkani, Riadin, Ni'mah, & Setiawan, 2025).

This paper addresses that requirement directly. It reports the development and content validation of an instrument designed to measure, in the digital learning context, the three constructs set out above, organized according to an Input–Process–Output logic: holistic digital well-being as input, digital character development as process, and holistic digital literacy as output. The distinctive contribution of the instrument, and the focus of this paper, lies in its operationalization of the Islamic psychospiritual constructs that constitute the process dimension—*Istiqamah* and *Muraqabah* as the components of digital self-regulation, and *Adab* as the domain of ethical digital comportment—together with the spiritual dimension of well-being expressed through the servant's relationship with the Creator. The paper documents how an initial pool of 91 items was generated from theory and refined, through a rigorous two-round expert-panel process, into a final 38-item instrument, and it reports the content validity of that instrument as established through the Content Validity Index.

The study is guided by three questions. First, through what procedure can Islamic

psychospiritual constructs, alongside dimensions of digital well-being and literacy, be operationalized into a set of measurement items suitable for the digital learning context? Second, what is the content validity of the resulting instrument, as judged by an expert panel using the CVI approach? Third, what preliminary psychometric evidence regarding internal consistency and suitability for subsequent factor analysis is provided by the interim field administration of the instrument? In answering these questions, the paper makes two contributions. Substantively, it provides a content-validated instrument for constructs whose role in digital learning has rarely been operationalized for empirical investigation. Methodologically, it offers a transparent and replicable procedure for translating value-based constructs from an intellectual tradition into psychometrically defensible indicators—a procedure of relevance well beyond the specific case of Islamic constructs, to any effort to measure culturally or religiously grounded ideas in education.

2. Methodology

This study is design-and-development research, and the instrument reported here was developed and validated within the design-and-development phase of a larger research programme. It follows a scale-development sequence consistent with established best practice: construct specification and item generation from theory, followed by expert evaluation of content validity and refinement, and preliminary reliability assessment prior to full-scale field testing (Boateng, Neilands, Frongillo, Melgar-Quinonez, & Young, 2018; Churchill, 1979). The content-validation component, which is the focus of this paper, proceeded through two rounds of expert-panel evaluation. The first was directed towards improving a draft instrument; the second towards quantifying content validity formally by means of the CVI approach. Three sources of evidence inform the study: the quantitative ratings and written comments supplied by the experts across the two rounds, preliminary reliability evidence from an initial student pilot,

and interim psychometric evidence from the field administration now under way.

The initial item pool was drawn from a structured review of the literature and from the theoretical anchors on which the study rests: the eudaimonic model of psychological well-being (Ryff & Keyes, 1995); the higher objectives of Islamic law, together with the constructs of *Istiqamah*, *Muraqabah* and *Adab*; the social-cognitive theory of self-regulation (Zimmerman, 2000); the second extended unified theory of technology acceptance (Venkatesh, Thong, & Xu, 2012); the DigComp 2.2 framework (Vuorikari, Kluzer, & Punie, 2022); and the multidimensional model of digital literacy (Ng, 2012). The items were sourced from three different places: some items were drawn from existing instruments and modified, some rephrased in the digital learning context, while others were developed from scratch because there was no existing item that fitted the purpose. All these stages followed established practices in developing measurement instruments as well as guidelines for educational and psychological testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014; Kalkbrenner, 2021; Elangovan & Sundaravel, 2021). Altogether, 91 items grouped into eight elements emerged from the process.

Four experts were invited to assess the content validity of the instrument. The composition of the panel included expertise in the domain of the instrument as well as its methodology: educational/instructional technology, Islamic education and wellbeing, Malay language, and general studies/social science. Such a number is in line with the established practice where the minimum of three experts is recommended, the typical range lies between three and ten, and higher agreement among experts is expected if the panel is smaller (Lynn, 1986; Polit, Beck, & Owen, 2007). Two of the four experts took part in both stages. Their presence allowed the second round to judge how far the earlier feedback had been acted upon, while the remaining members brought an independent

view of the revised instrument. Convening a compact panel was a deliberate choice rather than a merely pragmatic one. Validating this instrument called for an uncommon combination of expertise in Islamic theology and educational values on the one hand and in digital educational measurement on the other, since every item had to be judged at once for its fidelity to a revelation-based construct and for its psychometric adequacy in the digital learning context. Experts qualified in both respects are scarce. A small and highly purposive panel, each of whose members held the requisite specialist knowledge, was therefore preferable to a larger one whose additional members would have lacked one domain or the other and whose judgements would accordingly have been less informative.

In the first round, which ran over four weeks, each expert rated the appropriateness of every element and item on a seven-point agreement scale ranging from strong disagreement to strong agreement and added written comments. Because the aim at this stage was improvement rather than final validation, agreement at the element level was dichotomised, with ratings of five and above coded as agreement, to yield an agreement index; the item-level comments were then subjected to thematic content analysis so that areas requiring revision could be identified. Combining the two forms of evidence allowed the panel both to confirm the adequacy of the underlying construct structure and to guide the detailed refinement of individual items.

Between the rounds the instrument underwent systematic refinement, guided throughout by the first-round feedback. Five principal actions were taken. The first reorganised the overall structure from eight elements into three constructs aligned with the Input–Process–Output framework. Second, items that were overlapping, repetitive, overly general, or misaligned with their construct were dropped or merged. Third, dual-barrelled items were split into single-concept items and potentially biased wording was removed. Fourth, the digital literacy construct was expanded from a narrow focus on technology acceptance to a comprehensive conception of holistic digital

literacy grounded in the five DigComp 2.2 domains and the multidimensional literacy model. Fifth, each item was explicitly contextualized to the online learning environment and mapped to a specific theoretical dimension, with the Islamic constructs operationalized through indicators explicitly tagged to *Istiqamah*, *Muraqabah*, *Adab*, and the preservation objectives of intellect and property. This refinement reduced the pool from 91 items to a final instrument of 38 items.

In the second round, the refined conceptual framework and the 38-item instrument were evaluated for content validity using the Content Validity Index (Polit & Beck, 2006; Zamanzadeh et al., 2015). Each expert rated the relevance of every item on a four-point scale; ratings of three or four were coded as relevant and ratings of one or two as not relevant. Two indices were computed. The item-level index (I-CVI) was calculated as the proportion of experts rating an item as relevant, and the scale-level index was calculated both as the average of the item-level indices (S-CVI/Ave) and as the proportion of items rated relevant by all experts (S-CVI/UA). For a panel of four experts, an I-CVI of 1.00 is required and an S-CVI/Ave of 0.90 or above indicates excellent content validity (Lynn, 1986; Polit, Beck, & Owen, 2007).

Prior to the content-validation rounds, the draft dual-response instrument had been subjected to an initial pilot study with 33 undergraduate students drawn from both online and conventional modes of study, a sample size within the range recommended for preliminary reliability assessment. Internal consistency was estimated using Cronbach's alpha, with the Spearman–Brown coefficient applied to two-item scales for which alpha is a less appropriate estimator (Eisinga, te Grotenhuis, & Pelzer, 2013). The pilot yielded coefficients ranging from 0.70 to 0.96, providing initial evidence of acceptable internal consistency and supporting the feasibility of the response format prior to expert validation and full-scale testing (Kalkbrenner, 2021).

The study was carried out under institutional research-ethics approval, which was obtained before any data were collected. Participation was voluntary throughout and rested on informed consent. The 33 undergraduate students in the initial reliability pilot, together with the respondents to the subsequent field administration, were told the purpose of the study, the voluntary and anonymous character of their participation, and their right to withdraw at any point without penalty; each gave consent before completing the instrument. The four members of the expert panel consented on the same basis, having been informed of the purpose and procedures of the content-validation exercise and of the intended use of their anonymised ratings and written comments. All experts possessed doctoral qualifications and between 10 and 25 years of professional experience in their respective fields. Their identities are not disclosed. No personally identifiable information is reported, and all data were stored securely and used solely for the research purposes described here.

Preliminary Field Evaluation

Once content validation was complete, preliminary field evidence was examined using an interim dataset drawn from the main data-collection phase, which is still under way. At the cut-off, 274 responses had been received. Six respondents who did not give consent were excluded, as were 66 responses that failed the embedded attention check, leaving 202 eligible cases. The attention check instructed respondents to select a predetermined response option, and failure to do so served as the prespecified exclusion criterion. Of the eligible cases, 195 were complete across the 38

substantive scored items. The instrument also carried four response-quality indicators, comprising the explicit attention check and three further response-quality items, none of which entered construct scoring.

The interim analysis was confined by design to preliminary psychometric diagnostics rather than to full construct validation. Internal consistency was examined through Cronbach's alpha, and corrected item-total associations were inspected in order to identify items or subdimensions that might prove weak. Sampling adequacy for later examination of the latent structure was assessed by means of the Kaiser–Meyer–Olkin (KMO) measure and Bartlett's test of sphericity. Confirmatory factor analysis, the assessment of convergent and discriminant validity, and structural-model testing were all reserved for the completed field dataset.

3. Results

Table 1 summarises the composition of the expert panel across the two rounds. In the first round every one of the eight elements of the draft instrument attained a unanimous agreement index of 1.00, with element means ranging from 5.50 to 6.25 on the seven-point scale. The panel therefore judged the underlying construct structure to be sound, which warranted retaining the substantive content of the instrument through the refinement that followed. Agreement of this strength is striking, though it should be read in the light of what the round was for: the high element-level agreement established that the constructs were appropriate, while the detailed work of improvement was directed by the qualitative feedback rather than by the quantitative index.

Table 1. Expert Panel Composition Across the Two Validation Rounds

Expertise domain	Round 1	Round 2
Educational and instructional technology	✓	✓
Islamic education and well-being	✓	✓
Malay language	✓	✓

General studies (MPU) and social science	✓	✓
Panel size	4	4
Overlapping members retained	—	2

Source: Authors' compilation of expert panel records

Note. Continuity of two panel members across rounds enabled assessment of whether first-round feedback had been addressed while retaining independent judgement of the refined instrument.

The thematic analysis of the first-round comments identified eight recurring themes requiring attention, summarized in Table 2. These concerned the presence of dual-barrelled items that combined two ideas in a single

statement, sentences of excessive length, terms with the potential to introduce bias, overlap of meaning between items, and, most substantively, a tendency in the digital literacy content to emphasize the acceptance of technology at the expense of a fuller conception of digital competence. These themes set the agenda for the refinement described in the method, and their resolution can be traced in the reduction of the item pool and the restructuring of the constructs.

Table 2. Themes From Round 1 Expert Feedback and Corresponding Refinement Actions

Feedback theme	Refinement action
Dual-barrelled items	Split into single-concept items
Overly long sentences	Shortened and simplified
Potentially biased terms	Reworded for neutrality
Overlap of meaning between items	Merged or removed redundant items
Overly general or vague items	Dropped or rewritten for specificity
Digital literacy over-focused on technology acceptance	Expanded to five DigComp 2.2 domains and Ng's model
Insufficient contextualization	Each item anchored to the online learning context
Loose theoretical mapping	Each item mapped to a specific dimension and Islamic construct

Source: Thematic analysis of Round 1 expert comments by the authors

Refinement brought the initial pool of 91 items down to a final instrument of 38, with 53 items either dropped or merged in the light of expert feedback and further conceptual reflection. Table 3 sets out the instrument that resulted, organised into three constructs aligned with the Input–Process–Output framework. The input construct, holistic digital well-being, comprises five dimensions that operationalise the balanced development of the whole person. The process

construct, digital character development, comprises two dimensions, digital self-regulation and digital *Adab*. The output construct, holistic digital literacy, comprises five dimensions corresponding to the DigComp 2.2 competence areas. Each dimension is represented by three or four items, which supplies multiple indicators while keeping the measurement parsimonious (Hair et al., 2021).

Table 3. Structure of the Final 38-Item Instrument

Construct (IPO role)	Dimensions	Islamic operationalization	Item code
Holistic Digital Well-Being (Input)	Physical, emotional, spiritual, intellectual, social	Hablum Minallah / tuma'ninah; 'Afiyah; Hifz al-Nafs	KHD
Digital Character Development (Process)	Digital self-regulation	<i>Istiqamah</i> ; <i>Muraqabah</i> ; Hifz al-'Aql	RKD
Digital Character Development (Process)	Digital <i>Adab</i>	<i>Adab</i> ; netiquette; responsible conduct	ADB
Holistic Digital Literacy (Output)	Information/data, communication, content creation, safety, problem solving	Hifz al-Mal (data and asset protection); critical evaluation (tabayyun)	LDH

Source: Authors' construction of the refined instrument

Note. KHD = Kesejahteraan Holistik Digital; RKD = Regulasi Kendiri Digital; ADB = *Adab* Digital; LDH = Literasi Digital Holistik. The scoring instrument comprises 38 substantive construct items. A further four response-quality indicators, namely one explicit attention check and three response-quality items, were administered but excluded from construct scoring.

What matters about this structure, for the concern at the heart of the paper, is that every Islamic construct is represented by specific and contextualised indicators rather than by general or normative statements. Within the self-regulation dimension, *Istiqamah* and *Muraqabah* are operationalised as concrete capacities exercised in digital study. *Adab* is operationalised as observable practices of courteous and responsible interaction. The spiritual dimension of well-being is operationalised through the learner's relationship with the Creator as a source of tranquillity. In each case the construct becomes measurable without being reduced to behaviour stripped of its inner orientation, since

the indicators are framed to capture the disposition as it shows itself in the learner's digital conduct.

The second-round evaluation established the content validity of the refined framework and instrument, and the results are summarised in Table 4. Each of the three dimensions of the conceptual framework (input, process and output) obtained an item-level CVI of 1.00, which indicates unanimous expert agreement on the appropriateness of the model's structure. Every one of the 38 items obtained an I-CVI of 1.00 as well, and the scale-level indices stood correspondingly at their maximum, with S-CVI/Ave and S-CVI/UA both equal to 1.00. These values meet the threshold of 1.00 required for item-level agreement in a four-expert panel and surpass the 0.90 criterion for excellent content validity at the scale level (Lynn, 1986; Polit & Beck, 2006). Although unanimous agreement was observed, these findings should be interpreted as evidence of content validity rather than confirmation of construct validity.

Table 4. Content Validity Index Results, round 2

Object of validation	I-CVI	S-CVI/Ave	S-CVI/UA
IPO conceptual framework (3 dimensions)	1.00	1.00	1.00
Final instrument (38 items)	1.00	1.00	1.00

Source: Authors' computation from Round 2 expert ratings

Note. I-CVI = item-level content validity index; S-CVI/Ave = scale-level index, averaging method; S-CVI/UA = scale-level index, universal agreement method. For a four-expert panel an I-CVI of 1.00 is required, while an S-CVI/Ave of .90 or above indicates excellent content validity.

Of particular importance to the argument advanced here, the unanimity extended specifically to the dimensions in which the Islamic constructs are operationalised, namely digital self-regulation through *Istiqamah* and *Muraqabah*, and digital *Adab*. The panel brought together expertise in Islamic education and in educational measurement, and its agreement therefore affirms that these constructs had been rendered into indicators whose relevance can be defended on methodological grounds.

Preliminary Psychometric Evidence

The interim field dataset offers preliminary evidence on how the content-validated instrument performs empirically. Of the 202 eligible responses, 195 were complete across all 38 scored items. Internal consistency proved satisfactory to strong at the construct level, as Table 5 shows: holistic digital well-being yielded $\alpha = .925$, digital self-regulation $\alpha = .865$, digital *Adab* $\alpha = .790$, and holistic digital literacy $\alpha = .925$. The 38-item correlation matrix showed excellent sampling adequacy ($KMO = .930$), and Bartlett's test of sphericity was statistically significant, $\chi^2(703) = 5834.78$, $p < .001$, which indicates that the pattern of inter-item correlations is suitable for later examination of the latent structure.

Table 5. Preliminary Psychometric Evidence from the Interim Field Sample

Construct	No. of items	Cronbach's α
Holistic Digital Well-Being	15	.925
Digital Self-Regulation	4	.865
Digital <i>Adab</i>	4	.790
Holistic Digital Literacy	15	.925

Source: Authors' analysis of the interim field dataset

Note. Eligible interim sample $n = 202$; complete cases across the 38 scored items $n = 195$. $KMO = .930$; Bartlett's test of sphericity, $\chi^2(703) = 5834.78$, $p < .001$. These results represent preliminary psychometric evidence rather than completed construct validation. *Istiqamah* and *Muraqabah* were operationalised as complementary psychospiritual components within the Digital Self-Regulation dimension and were not scored as independent subscales.

Reliability at the broader construct level was satisfactory. Examination of the five digital-literacy subdimensions, however, revealed weaker internal consistency for the three-item safety dimension ($\alpha = .443$). Item-level diagnostics traced this to the item on regular password changes, which associated comparatively weakly with the remaining safety indicators. Since the present analysis rests on an

interim dataset, the item was retained pending evaluation against the completed field sample rather than removed post hoc. The finding illustrates how content relevance established by experts and empirical item performance can diverge, and it marks out one aspect of the instrument that calls for closer psychometric scrutiny.

4. Discussion

The central achievement reported here is the operationalisation of selected Islamic psychospiritual constructs into a content-valid measurement instrument situated specifically within digital education. It matters because the want of such operationalisation has long obstructed the empirical study of these ideas. For as long as *Istiqamah*, *Muraqabah* and *Adab* remain normative ideals affirmed in the abstract, claims about their role in learning cannot be tested, and

the conceptual resources of the Islamic tradition cannot be brought into productive contact with the empirical methods of educational research. By rendering them into indicators that a panel of experts, including specialists in Islamic education and in measurement alike, judged unanimously to be relevant, the instrument allows them to be examined through empirical measurement approaches that have more commonly been applied to established psychological constructs. The contribution is not confined to the research programme within which the instrument was developed; it extends to the wider project of integrating value-based and religiously grounded constructs into educational measurement (Zafar & Malik, 2025; Aycan, 2024).

The procedure by which this was accomplished is itself a contribution and merits reflection. Its two-round structure, in which the first round sought improvement through combined quantitative and qualitative evaluation while the second quantified content validity formally, proved well suited to constructs of this difficulty. The first round confirmed the adequacy of the construct structure and, through thematic analysis of expert comments, brought to the surface the specific deficiencies that item generation from theory had left in the draft. The refinement that followed addressed these systematically, and the second round then quantified the content validity of the revised instrument. That sequence embodies a principle stressed throughout the scale-development literature: content validity is established not by a single evaluative act but by an iterative process of expert judgement and revision (Almanasreh, Moles, & Chen, 2019; Boateng et al., 2018). The particular difficulty of psychospiritual constructs, namely the risk of capturing outward behaviour while missing the inner orientation that gives it meaning, was met by contextualising every item explicitly to the digital learning environment and by mapping each to a specific dimension of its construct. The indicators therefore capture the disposition as it manifests in conduct rather than the conduct alone.

The results also speak to a methodological concern that attends any report of unanimous

expert agreement, namely that such agreement may reflect insufficiently critical evaluation rather than genuine validity. Two features of the present study address this concern. The first is the transparency of the refinement process: the unanimous second-round agreement is intelligible precisely because the instrument had been extensively revised in response to substantive first-round criticism, so that the second round validated a product that had already survived critical scrutiny. The second is the independence of the expert judgements, each rendered without consultation among the panel. Together these features support the interpretation of the unanimous agreement as a defensible finding rather than an artefact of deference, though the study does not claim that content validity alone establishes the instrument's overall validity, which further requires the evidence of construct validity that subsequent psychometric evaluation must provide (Hair et al., 2021).

A further consideration is mathematical. With a four-member panel the item-level CVI can take only the values 1.00, 0.75, or lower, and because the accepted threshold for a panel of this size is an I-CVI of 1.00 (Polit, Beck, & Owen, 2007), a single dissenting judgement is sufficient to fail an item. That every retained item command unanimous agreement therefore follows from applying the stringent small panel criterion and is not evidence of an uncritical panel. Only items on which all four experts concurred could survive into the final instrument, and weaker items had in any case already been removed by the pre-validation refinement documented in Table 2. The same feature is nonetheless a limitation. With only four raters the index is coarse-grained and cannot register the finer gradations of agreement that a larger panel would afford, so construct validity remains to be established on field data.

The study also contributes to the more specific literature on the measurement of Islamic constructs. Recent years have brought validated scales for individual constructs such as *Istiqaamah*, along with systematic conceptual work comparing *Muraqabah* with related psychological

ideas (Zafar & Malik, 2025; Nurhakim & Rahman, 2025), and, within Islamic higher education, competence measures that embed Islamic ethical principles in the use of digital and AI technologies (Syamsudduha, Suwandi, & Tabrani ZA, 2026). The present instrument complements that work in two respects. It operationalises *Istiqamah*, *Muraqabah* and *Adab* specifically for the digital learning context, where their relevance is at once novel and pressing. It also brings them together with dimensions of well-being and literacy in a single content-validated instrument, so that what can be studied is not the individual constructs in isolation but their relationships within a coherent model. In doing so it answers a call now heard with growing frequency in the study of digital education for measurement tools that reflect the whole person rather than technical competence alone (Weber, Kammerl, & Schiefner-Rohs, 2025; Georgopoulou, Troussas, Krouska, & Sgouropoulou, 2025).

The preliminary field evidence offers an important complement to the expert-based findings. Content validity establishes that the indicators are conceptually relevant to the constructs they are meant to represent. Internal consistency and sampling adequacy answer a different question, namely how those indicators actually perform and whether the correlation matrix is suitable for subsequent factor analysis. That construct-level reliability proved satisfactory to strong, and that the KMO in the interim sample was excellent, together strengthen the case for proceeding to full construct validation. The weaker reliability of the safety subdimension, by contrast, cautions against treating unanimous expert agreement as sufficient evidence of psychometric adequacy. This divergence reinforces the sequential logic of instrument development: theoretically grounded operationalisation and expert judgement lay the content foundation, while field-based analysis determines whether the proposed measurement structure holds empirically.

Several limitations qualify these findings. Although the interim field analysis extends the evidence beyond expert judgement by

demonstrating preliminary internal consistency and suitability for subsequent factor analysis, it does not establish the instrument's construct validity. The field dataset was incomplete at the prespecified analysis cut-off, and the present study therefore deliberately refrained from confirmatory factor analysis, convergent and discriminant validity assessment, or structural-model testing. These analyses require evaluation using the completed dataset. The weaker preliminary reliability of the safety subdimension also warrants further item-level investigation. In addition, the expert panel, while appropriately composed for content validation, was small, and the instrument was developed within a specific cultural and institutional context. Replication and cross-contextual validation will therefore be necessary before broader generalizability can be claimed.

5. Conclusion

This paper has reported the development, content validation, and preliminary psychometric evaluation of a 38-item instrument that operationalizes Islamic psychospiritual constructs—*Istiqamah*, *Muraqabah*, and *Adab*—together with dimensions of holistic digital well-being and holistic digital literacy, for the digital learning context. Through a two-round expert-panel process, an initial pool of 91 items was refined into a final instrument whose conceptual framework and every item obtained a content validity index of 1.00, with scale-level indices at their maximum. The study demonstrates that constructs frequently invoked as normative ideals, yet still under-operationalised in digital education research, can be translated into measurable indicators with defensible content validity. What such a translation requires is careful conceptual specification, systematic expert evaluation and iterative refinement. The consequences for educational research and the practice of Islamic education are twofold. For science, the instrument supplies an operational basis on which the roles of *Istiqamah*, *Muraqabah* and *Adab* can be examined in hypothesis-driven research on digital education, so that their contribution to learning may be investigated empirically rather than assumed. For praxis, it

equips institutions of Islamic education with a means of assessing the psychospiritual formation of learners in digital environments, and thus of designing and evaluating interventions on evidence rather than on assumption. Taken together, the expert-validation results and the preliminary field evidence establish a defensible foundation for the next stage of instrument evaluation. Full-sample psychometric testing is now required to examine the latent structure, convergent and discriminant validity, and ultimately the proposed relationships among holistic digital well-being, digital character development, and holistic digital literacy.

6. Acknowledgements

The authors would like to express their sincere appreciation to the four expert panel members, the students who participated in the pilot study, and the respondents involved in the field administration for their time, valuable contributions, and thoughtful judgments. The authors also gratefully acknowledge the support provided by UNITAR International University throughout the conduct of this research. In addition, the authors extend their sincere thanks to the researchers from Universiti Kebangsaan Malaysia (UKM) for their academic guidance and expert advice, which contributed significantly to the successful implementation of the exploratory phase of this study.

References

- Al-Khatib, T. (2023). Netiquette rules in online learning through the lens of the digital citizenship scale in the post-corona era. *Journal of Information, Communication and Ethics in Society*, 21(2), 181–201. <https://doi.org/10.1108/JICES-826-839-0089>
- 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>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Aycan, S. (2024). Construction of an Islamically-integrated psychological well-being model. *Journal of Muslim Mental Health*, 18(1), 46–61. <https://doi.org/10.3998/jmmh.6026>
- Bagis, F., Adawiyah, W. R., Purnomo, R., & Sudjadi, A. (2024). Exploring Islamic spiritual well-being: Conceptualization and validation of a measurement scale. *Islamic Guidance and Counseling Journal*, 7(2). <https://doi.org/10.25217/0020247473500>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>

7. Author Declaration

The authors confirm that this manuscript is original work, that it has not been published previously, and that it is not under consideration for publication elsewhere. The authors received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors for the research, authorship, or publication of this article, and declare no potential conflicts of interest. The study received institutional research-ethics approval, and informed consent was obtained from all participants, as described in the method. The data that support the findings of this study are available from the corresponding author upon reasonable request; they are not publicly available owing to privacy and ethical restrictions.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work the author(s) used Claude (Anthropic) in order to improve the readability and language of the manuscript. After using these tools and services, the author(s) reviewed and verified the content and all reported statistics, edited them as needed, and take(s) full responsibility for the content of the published article.

- Bulkani, B., Riadin, A., Ni'mah, N., & Setiawan, M. A. (2025). Impact of holistic learning models on character development: A systematic review. *Obrazovanie i Nauka*, 27(5), 111–141. <https://doi.org/10.17853/826-839-826-839-826-839>
- Chaudhary, M. Y. (2020). Initial considerations for Islamic digital ethics. *Philosophy and Technology*, 33(4), 639–657. <https://doi.org/10.1007/s1826-839-0826-839>
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64–73. <https://doi.org/10.1177/002224377901600110>
- Eisinga, R., te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? *International Journal of Public Health*, 58(4), 637–642. <https://doi.org/10.1007/s0826-839-826-839>
- Elangovan, N., & Sundaravel, E. (2021). Method of preparing a document for survey instrument validation. *MethodsX*, 8, 101326. <https://doi.org/10.1016/j.mex.2021.101326>
- Georgopoulou, M. S., Troussas, C., Krouska, A., & Sgouropoulou, C. (2025). Digital literacy in higher education: Examining university students' competence in online information practices. *Computers*, 14(12), 528. <https://doi.org/10.3390/computers14120528>
- Habib, Z. (2025). Ethics of artificial intelligence in Maqasid al-Shari'a's perspective. *KARSA: Journal of Social and Islamic Culture*, 33(1), 105–134. <https://doi.org/10.19105/karsa.v33i1.19617>
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.
- Kalkbrenner, M. T. (2021). Alpha, omega, and H internal consistency reliability estimates: Reviewing these options and when to use them. *Counseling Outcome Research and Evaluation*, 14(1), 77–88. <https://doi.org/10.1080/21501378.2021.1940118>
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. <https://doi.org/10.1097/0000826-8391826-8397>
- Md Razali, S., & Mohd Fithri, K. (2024). Psychospiritual well-being through understanding i'jaz al-Qur'an. *Quranica*, 16(1), 83–105. <https://doi.org/10.22452/quranica.vol16no1.4>
- Nair, S. D., & Nesamany, S. S. (2021). Teacher perception of online teaching approaches during the pandemic COVID-19. *Asian Journal of Social Science Research*, 3(2), 20–38. <https://doi.org/10.5281/zenodo.8063275>
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078. <https://doi.org/10.1016/j.compedu.2012.04.016>
- Nurhakim, A., & Rahman, A. (2025). The lhan Hadith as the spiritual foundation of *Muraqabah*: A comparative study with contemporary mindfulness. *Afkar*, 27(2), 427–456. <https://doi.org/10.22452/afkar.vol27no2.13>
- 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>
- Rifin, R., Sidik, N. B., Abdul Rahman, N. H., Mohd Hussain, M., & Mokhtar, S. M. A. (2025). The influence of self-regulated strategies on components in motivational beliefs. *Asian Journal of Social Science Research*, 7(1), 146–160. <https://doi.org/10.5281/zenodo.15804446>
- Ryff, C. D., & Keyes, C. L. M. (1995). The structure of psychological well-being revisited. *Journal of Personality and Social Psychology*, 69(4), 719–727. <https://doi.org/10.1037/826-839.69.4.719>
- Shahwar, D., & Dhar, R. L. (2024). Unravelling the mysteries of cyber incivility: A systematic review and research agenda. *International Journal of Conflict Management*, 35(4), 658–683. <https://doi.org/10.1108/IJCMA-826-839-0115>
- Syamsudduha, S., Suwandi, S., & Tabrani ZA. (2026). Islamic AI literacy as responsible knowledge practice: Development and validation of AILS-IIUS in Indonesian Islamic higher education. *Jurnal Pendidikan Islam*, 12(1), 187–201. <https://doi.org/10.15575/jpi.v12i1.54254>

- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. <https://doi.org/10.2307/41410412>
- Vuorikari, R., Kluzer, S., & Punie, Y. (2022). *DigComp 2.2: The Digital Competence Framework for Citizens—With new examples of knowledge, skills and attitudes*. Publications Office of the European Union.
- Weber, P. M., Kammerl, R., & Schiefner-Rohs, M. (2025). What does digital well-being mean for school development? A theoretical review with perspectives on digital inequality. *Education Sciences*, 15(8), 948. <https://doi.org/10.3390/educsci15080948>
- 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>
- Zafar, J., & Malik, N. I. (2025). Development and validation of the Istiqamah Scale for Pakistani Muslim adults. *Journal of Religion and Health*, 64(4), 3167–3191. <https://doi.org/10.1007/s1826-839-02301-x>
- Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation of 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>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press. <https://doi.org/10.1016/B826-8390826-839/5826-839>