Developing the Islamic scale of wisdom – general version (ISW-GV)

Ahmad Rusdi[1*], Fani Eka Nurtjahjo[2], Nyda Afsari[3], Ira Mayang Sari[4], Resha Karina Puteri[5], Anis Syifa Nuraini[6], Audi Vidya Putri Amirza[7]

[1,2,3,4,5,6,7,8]Faculty of Psychology, Universitas Islam Indonesia, Indonesia
Corresponding Author*: ahmad_rusdi@uii.ac.id

Article Info	ABSTRACT

The lack of a culturally grounded instrument to assess wisdom (*ḥikmah*) in general Muslim populations necessitates the development of a contextually relevant scale. This study introduces the Islamic Scale of Wisdom – General Version (ISW-GV), constructed from Ibn Miskawaih's conceptualization of *ḥikmah* as a balanced cognitive-reflective virtue, addressing the limitations of Western-based wisdom measurements that tend to emphasize cognitive and experiential aspects while overlooking the moral and virtuous dimensions central to Islamic thought. Research was conducted across 857 respondents through Exploratory and Confirmatory Factor Analysis, and correlational tests. Results supported a stable four-factor structure (problem solving, learning, reflecting, and understanding) with satisfactory model fit indices. Convergent evidence was confirmed through positive correlations with HS, ISW-AV, and 3D-WS, while discriminant evidence was demonstrated via weak associations with GPA and social desirability. Criterion-related analysis showed that wisdom correlated positively with subjective happiness among adult workers but not among students, suggesting developmental variation in the function of *ḥikmah*. These outcomes support ISW-GV as a valid and context-sensitive instrument for capturing wisdom across everyday settings. However, response patterns indicate that the current bipolar item format (1-2-3-2-1) may blur distinctions between *ḥikmah*, rashness, and stupidity. Future refinements using rank-order Situational Judgment Test formats may enhance its assesment precision in mapping ethical reasoning across value orientations.

INTRODUCTION

Moral virtues are central in the life of a Muslim. They serve as the foundation of goodness and well-being (Bergsma & Ardelt, 2012; Nabi, 2024). Among the core virtues, *ḥikmah* (wisdom) occupies a pivotal role as the intellectual disposition that directs behavior. In *Tahdhīb al-Akhlāk*, Ibn Miskawaiḥ (2011) situates *ḥikmah* alongside *shajāʿah* (courage), *ʿiffah* (temperance), and *sakhāʾ* (generosity) as the four fundamental virtues. When balanced, these generate justice. *Ḥikmah* reflects the golden mean in the use of reason. It avoids both rashness (*al-safāh*) and foolishness (*al-balāh*), cultivating proportional judgment and action (Othman et al., 2021). A lack of intellectual control, by contrast, may result in cunning or deceit. Hence, *ḥikmah* extends beyond intellectual ability. It represents the integration of sound reasoning, moral awareness, and behavior aligned with Islamic values (Khan, 2025).

In contemporary psychology, wisdom is defined as a multidimensional personality trait comprising three core aspects: (1) cognitive, reflective, and affective traits (Ardelt, 2003); (2)

adaptive problem-solving (Krieger & Greiff, 2023); and (3) integration of intelligence, creativity, and ethics in decision-making (Sternberg et al., 2021). In positive psychology, wisdom is classified as a core character strength (Peterson & Seligman, 2004). These perspectives highlight the synthesis of knowledge, experience, and emotion regulation (Kaymak et al., 2025). In contrast, Ibn Miskawaih defines *ḥikmah* (wisdom) as the proportional use of intellect, cultivated through education and training rather than acquired solely from experience. Consequently, wisdom is rooted in education (*tarbiyyah*) and cultural cultivation (Alfaiz et al., 2025; Mahpur et al., 2023).

Scale development in Western psychology has produced instruments such as the Three-Dimensional Wisdom Scale (3D-WS; Ardelt, 2003), which has been adapted for Indonesia (Indati et al., 2019; Raharja & Indati, 2019). Its cultural fit for Indonesian Muslim populations remains untested. In parallel, Islamic scholarship provides distinct conceptualizations of wisdom. Efforts to develop Islamic-based instruments include Utami's (2020) al-Ghazālī-inspired scale, which has undergone preliminary Rasch validation among students, though its factorial structure is undetermined. The Islamic Scale of Wisdom – Academic Version (ISW-AV; Nurtjahjo et al., 2021; Nurtjahjo & Rusdi, 2018) is designated for higher education. Currently, no psychometrically validated instrument exists to measure *ḥikmah* as wisdom within Islamic ethics for general population.

This study addresses the gap by developing and validating the Islamic Scale of Wisdom – General Version (ISW-GV), precisely grounded in Ibn Miskawaih's framework. Here, *ḥikmah* comprises seven defined aspects: (1) *al-dhakāʾ* (deductive intelligence), (2) *al-taʿaqqul* (objective reasoning), (3) *suhūlat al-taʿallum* (ease of learning), (4) *surʿat al-fahm* (accuracy of comprehension), (5) *al-dhikr* (relevant memory), (6) *jawdah al-zihn* (quality of reflection), and (7) *ṣafāʾ al-zihn* (clarity of thought). Each represents the golden mean, avoiding both deficiency (*al-balāh*) and excess (*al-safāh*). Thus, *ḥikmah* is a balanced cognitive-reflective capacity that underpins rational and morally attuned behavior.

Ibn Miskawaih (2011) in *tahdhīb al-akhlāq* describes seven key derivations of *ḥikmah*. Each has its own balanced form and corresponding extremes. First, *al-dzakāʾ* refers to the ability to draw accurate conclusions quickly and clearly. Its deficiency, *al-balādah*, reflects an untrained or sluggish intellect. Its excess, *al-khabath*, reflects sharp reasoning used manipulatively or immorally. Balanced *al-dzakāʾ* represents intellectual clarity guided by ethical purpose (Ibn Miskawaih, 2011). Second, *al-taʿaqqul* captures the capacity for objective thinking—evaluating facts in accordance with reality. Its deficient form, *al-qaṣur bi al-naẓar*, leads to narrow, shallow, or weak analysis. Its excessive form, *al-dhahāb bi al-naẓar*, produces overanalysis that overwhelms clarity. The midpoint describes proportionate, well-grounded reasoning (Ibn Miskawaih, 2011). Third, *suhūlah al-taʿallum* refers to ease in learning theoretical or conceptual knowledge. The deficient form, *al-taʿaṣṣub*, denotes difficulty in understanding abstraction and rigid adherence to preexisting assumptions. The excessive form, *al-mubādarah*, denotes overly rapid but superficial understanding. The balanced form accurate comprehension in learning (Ibn Miskawaih, 2011). Fourth, *surʾah al-fahm* involves the ability to grasp straightforward information quickly and correctly. Its deficiency, *al-ibṭāʾ*, indicates slowness or difficulty in comprehending relationships. Its excess, *ikhtiṭāf*, denotes overly fast but inaccurate comprehension. Wisdom lies in maintaining both speed and correctness (Ibn Miskawaih, 2011). Fifth, *al-dhikr* denotes memory—the ability to retain knowledge or past events in a stable and accurate form. *al-Nisyān* represents the deficient extreme, where important matters are neglected or forgotten. The extreme, *al-ʾināyah*, is where attention is absorbed by details that should not be remembered. Balanced memory prioritizes what is valuable and beneficial (Ibn Miskawaih, 2011). Sixth, *jaudah*

al-zihn concerns reflective capacity: the ability to reflect on past experiences in proportion and draw meaningful lessons. Its deficiency manifests as limited or absent reflection. Its excess manifests as overreflection, unproductive rumination, or regret that hinders action. The midpoint represents mature and proportionate reflection (Ibn Miskawaih, 2011). Seventh, *ṣafā' al-zihn* refers to mental clarity arising from emotional stability. The deficient extreme, *ẓulmah al-nafs*, reflects an emotionally clouded inner state that impedes clear understanding. The excessive extreme, *iltihāb al-nafs*, reflects heightened emotional agitation—such as anger or resentment—that disrupts clarity. Balanced *ṣafā' al-zihn* represents a calm and receptive internal state enabling sound conclusions (Ibn Miskawaih, 2011).

While prior studies have introduced scales inspired by concepts of *tahdhīb al-akhlāq* Nurtjahjo et al., 2021; Nurtjahjo & Rusdi, 2018; Utami, 2020), many have overlooked validation procedures. These include content validity, response process validity, and confirmatory factor analysis. Moreover, adaptations of Western scales (Indati et al., 2019; Raharja & Indati, 2019; Vaisi et al., 2025) have not been examined for cultural congruence with Muslim populations. The ISW-GV offers a novel contribution by extending the scope of the ISW-AV to general contexts and by adhering to rigorous psychometric procedures. This study, therefore, advances both the theoretical understanding of *ḥikmah* and the empirical tools available for its measurement, ensuring relevance to Islamic ethics while maintaining methodological robustness. This study systematically developed and validated the ISW-GV through item construction, factor analysis, and validity assessment. The goal is to develop a robust, relevant, and empirically sound measure of *ḥikmah* for Muslim populations.

METHODS

This study employed a multi-stage process of instrument development and validation, including scale construction, content and face validity assessment, exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and evaluation of convergent and criterion-related validity. This study involved two rounds of data collection. The first study, used for the EFA, included 319 participants. The second study, used for the CFA, criterion-related validity, convergent, and discriminant validity, involved 538 participants. The following is a more detailed description of the demographic profile from the second study.

Table 1

Respondent Demographic Profile

No	Baseline Characteristic	Students		College Students		Workers		N (%)
		n	*%*	*n*	*%*	*n*	*%*	
1	Sex							
	Male	51	9,5%	19	3,5%	102	19%	172 (31.9%)
	Female	143	26,6%	88	16,4%	135	25%	366 (68.1%)
2	Age Distribution							
	16-19 years old	194	36%	34	6,3%	0	-	228 (42.4%)
	20-40 years old	0	-	73	13,6%	159	29,6%	232 (43.1%)
	41-65 years old	0	-	0	-	77	14,3%	77 (14.3%)
	Above 65 years old	0	-	0	-	1	0,2%	1 (0.001%)

A total of 857 participants were recruited across two stages: EFA (n = 319) and CFA with criterion-related convergence and discriminant validity (n = 538). Eligible participants were male or

female, aged 16 years or older, and had no reading or writing impairments. Of these, 265 (31%) were male, and 592 (69%) were female. Age distribution was as follows: 300 participants (35%) were 16-19 years old, 419 (48.9%) were 20-40 years old, and 137 (16%) were 41-65 years old. Of these, 265 (31%) were male, and 592 (69%) were female.

The Hikmah Subscale–Noble Character (HS) developed by Utami (2020) was employed. This 15-item scale uses a four-point Likert format ranging from 1 (strongly disagree) to 4 (strongly agree). It assesses two behavioral indicators: acting in mutually beneficial ways without harming oneself or others through deliberation and consensus, and achieving balance through wise conduct. Reported item reliability was 0.95, with Cronbach's alpha of 0.84. The HS was included for convergent validity testing, given its conceptual alignment with ḥikmah.

The Islamic Scale of Wisdom – Academic Version (ISW-AV) (Nurtjahjo et al., 2021; Nurtjahjo & Rusdi, 2018) comprises 26 items across six aspects: comprehension, control, problem-solving, scholastic, fast reasoning, and stability. The items were rated on a six-point Likert scale. Overall reliability was 0.914, with subscale reliabilities ranging from 0.643 to 0.906. ISW-AV was administered to assess convergent validity, as it reflects constructs of wisdom in Islamic contexts.

The Three Dimensional Wisdom Scale (3D-WS; Ardelt, 2003), translated into Indonesian (Indati et al., 2019; Raharja & Indati, 2019), contains 12 items across cognitive, reflective, and affective dimensions. Cronbach's alpha coefficients were 0.79, 0.77, and 0.72 for the respective dimensions. Items were rated on a five-point scale (1 = strongly agree, 5 = strongly disagree). Prior exploratory and confirmatory analyses support its structure. The 3D-WS was used to further assess convergent validity. To detect potential response bias, the Brief Social Desirability Scale (BSDS) (Haghighat, 2007) was used. This four-item scale ($\alpha = 0.60$) ranges from 0 to 4, with higher scores indicating greater social desirability. It was utilized for discriminant validity testing, as socially desirable responding could threaten the validity of the ḥikmah scale.

Cumulative Grade Point Average (GPA) was included as an indicator of academic achievement and factual-analytical knowledge. GPA reflects mastery of domain-specific content, logical reasoning, and technical competence fostered through formal education (Liu et al., 2025). In contrast, wisdom (ḥikmah) emphasizes proportional reasoning, moral discernment, and the integration of cognitive, reflective, and affective capacities directed toward life values (Jifa, 2013). Thus, while GPA captures performance in structured academic contexts, wisdom transcends academic success by guiding appropriate judgment and ethical conduct. GPA was therefore used to examine discriminant validity, given its conceptual distinction from the broader construct of wisdom.

The Subjective Happiness Scale (SHS) (Lyubomirsky & Lepper, 1999), previously translated and used in Indonesia (Rusdi et al., 2022), is a four-item measure of global happiness ($\alpha = 0.704$). Selection was based on evidence linking wisdom to happiness (Cheng et al., 2025; Omar et al., 2021; Siah, 2025). The SHS was used to assess criterion-related validity. Validity evidence based on test content was examined using the Content Validity Index (CVI) and Content Validity Ratio (Ayre & Scally, 2014; Shi et al., 2012; Wilson et al., 2012). Face validity was assessed through the Face Validity Index (FVI) at both item (I-FVI) and scale (S-FVI) levels (Yusoff, 2019). FVI values above 0.80 were considered acceptable. Interviews were conducted with a subset of respondents to further ensure item clarity, with expert judgment informing final item selection (Gültürk, 2024).

EFA was conducted in SPSS 22.0 using principal component analysis with promax rotation (Ahmed & Maruod, 2025). Sampling adequacy was assessed via the Kaiser–Meyer–Olkin statistic >0.70 (Lloret et al., 2017) and Bartlett's test of sphericity $p < 0.05$ (Sigudla & Maritz, 2023). Factors with eigenvalues ≥1 were retained (Alavi et al., 2020). Items with factor loadings ≥0.32 were

considered acceptable (Güvendir & Özkan, 2022), while those with cross-loadings ≥0.32 were reviewed for removal.

CFA was performed using AMOS to test the measurement model. Model fit was evaluated with χ^2, RMSEA, CFI, TLI, and GFI indices. Acceptable thresholds included RMSEA ≤0.05 (Garnier-Villarreal & and Jorgensen, 2025; Whittaker, 2012), CFI and TLI ≥0.90 (Hu & Bentler, 1999), and GFI >0.90. Discriminant validity was confirmed when inter-factor correlations were ≤0.85 (Cheung et al., 2023). RMSEA, CFI, and TLI are the primary model-fit indicators because these metrics are widely regarded as the most informative and theoretically appropriate indices for evaluating reflective measurement models (Hu & Bentler, 1999). RMSEA provides an absolute measure of model misfit while penalizing unnecessary complexity, whereas CFI and TLI evaluate incremental fit by comparing the proposed model with an independence baseline and adjusting for parsimony (Hu & Bentler, 1999). These three indices are consistently recommended in psychometric validation research and provide a stable, discriminating assessment of latent-structure quality. Other indices, such as SRMR, is not included because SRMR is less sensitive in multifactor oblique models and may yield overly optimistic results when factors are correlated; simulation studies by Hu and Bentler (1999) showed that combinations of RMSEA with CFI/TLI provide a more reliable basis for model evaluation. Criterion-related evidence was examined via Pearson (for interval data) and Spearman (for ordinal data) correlations. Among student participants, ISW-GV was correlated with HS, ISW-AV, 3D-WS, and SHS, with expected coefficients ranging from 0.30 to 0.60 (Ward et al., 2009). Correlations with BSDS and GPA were expected to be weaker (<0.50), reflecting discriminant evidence (Fiori & Antonakis, 2011).

RESULTS AND DISCUSSION

Scale Construction and Item Content

The development of the 24-item Hikmah Scale began with operationalizing the seven conceptual derivatives of *ḥikmah* identified from Ibn Miskawaih's framework into observable behavioral indicators. Each derivative, namely 1) *al-dzakā',* 2) *al-ta'aqqul,* 3) *suhūlah al-ta'allum*, 4) *sur'ah al-fahm*, 5) *al-dhikr*, 6) *jaudah al-zihn*, and 7) *ṣafā' al-zih,* was first translated into a set of concrete behaviors. These indicators were then converted into simple situational judgment–type items, yielding 24 items (three to four per derivative) that capture the full conceptual breadth of *ḥikmah*. The initial item pool was reviewed by two psychometric experts in a two-stage process: written expert review followed by an in-depth online cognitive interview to refine wording, clarity, and construct alignment.

The response format was designed according to the golden mean framework, using a five-option. The two opposite extreme options (A and E) were scored as 1, the two opposite "gradual" options (B and D) were scored as 2, and the midpoint option (C), representing the most proportional response, was scored as 3. This 1–2–3–2–1 scoring scheme was intended to capture balanced cognitive–reflective judgment and reduce acquiescence bias, as the optimal response was located at the midpoint rather than at an extreme (unlike conventional Likert scaling). Following item generation, two psychometric experts reviewed all items to evaluate indicator–item alignment, clarity, and the logical progression of response options. Their feedback informed revisions, particularly in ensuring that each option flowed coherently from one to the other. The revised version was then evaluated by 23 Subject Matter Experts (SMEs) through an online form. Each item was rated on a six-point scale (1 = very irrelevant to 6 = very relevant). Ratings of "fairly relevant," "relevant," and

"very relevant" were coded as essential (Ne = 1), while the lower three categories were treated as non-essential (Ne = 0). Content validity was assessed using the Content Validity Ratio (CVR) and Content Validity Index (CVI) (Ayre & Scally, 2014; Shi et al., 2012; Wilson et al., 2012).

Table 2

Validity Evidence Based on Test Content

Indicator	Aitem	Ne	I-CVR	I-CVI	Decission
al-Dzakā'	Item 1	20	0,739	0,869	Essential
	Item 2	21	0,826	0,913	Essential
	Item 3	17	0,478	0,739	Essential
	Item 4	19	0,652	0,826	Essential
al-Ta'aqqul	Item 5	19	0,652	0,826	Essential
	Item 6	20	0,739	0,869	Essential
	Item 7	23	1.000	1.000	Essential
Suhūlah al-ta'allum	Item 8	22	0,913	0,956	Essential
	Item 9	20	0,739	0,869	Essential
	Item 10	20	0,739	0,869	Essential
	Item 11	20	0,739	0,869	Essential
Sur'ah al-fahm	Item 12	22	0,913	0,956	Essential
	Item 13	21	0,826	0,913	Essential
	Item 14	21	0,826	0,913	Essential
al-dhikr	Item 15	23	1.000	1.000	Essential
	Item 16	23	1.000	1.000	Essential
	Item 17	22	0,913	0,956	Essential
Jaudah al-zihn	Item 18	23	1.000	1.000	Essential
	Item 19	23	1.000	1.000	Essential
	Item 20	22	0,913	0,956	Essential
Ṣafā' al-zihn	Item 21	23	1.000	1.000	Essential
	Item 22	23	1.000	1.000	Essential
	Item 23	23	1.000	1.000	Essential
	Item 24	23	1.000	1.000	Essential
	Overall Judgment	23	1.000	1.000	Essential

Table 2 presented the detailed CVR and CVI results. Critical values for N = 23 was CVR = 0.391 and recommendations for CVI ≥ 0.80 (Ayre & Scally, 2014), all 24 items surpassed the required thresholds. Specifically, CVR values ranged from 0.478 to 1.000, and CVI values ranged from 0.739 to 1.000. These results demonstrated strong content validity and a robust theoretical match between items and the construct of *ḥikmah*. Nevertheless, minor editorial refinements were made to improve item clarity. For instance, Item 3 initially contained two response options with overlapping meaning, which were revised to strengthen the gradation of responses without altering the underlying construct. The final version of the instrument was confirmed in consultation with psychometric experts, ensuring readiness for subsequent validation stages, including face validity (response process analysis), EFA, and CFA (Yusoff, 2019; Gültürk, 2024).

Validity Evidence on Response Processes

Response process evidence was examined through participant ratings of item clarity (FVI) and cognitive interviews. The face validity index (FVI) analysis demonstrated that the *ḥikmah* scale achieved satisfactory clarity and comprehensibility. Item-level FVI (I-FVI) ranged from 0.733 to

1.000, while the scale-level FVI (S-FVI/Ave) reached 0.904, exceeding the recommended threshold of 0.80 (Yusoff, 2019). Two items, Item 3 (I-FVI = 0.767) and Item 6 (I-FVI = 0.733), fell below the minimum cut-off and were therefore eliminated from the item pool.

Table 3

Face Validity Information

Aspects	Item	*Raters in agreement*	*Universal Agreement (UA)*	I-FVI
al-Dzakā'	Item 1	30	1	0.867
	Item 2	30	1	0.833
	Item 3	28	0	0.767
	Item 4	28	0	0.900
al-Ta'aqqul	Item 5	30	1	1.000
	Item 6	29	0	0.733
	Item 7	30	1	0.933
Suhūlah al-ta'allum	Item 8	30	1	0.900
	Item 9	30	1	0.967
	Item 10	29	0	0.933
	Item 11	30	1	0.967
Sur'ah al-fahm	Item 12	29	0	0.867
	Item 13	30	1	0.933
	Item 14	30	1	0.967
al-dhikr	Item 15	30	1	0.933
	Item 16	28	0	0.867
	Item 17	28	0	0.933
Jaudah al-zihn	Item 18	29	0	0.900
	Item 19	29	0	0.900
	Item 20	29	0	0.867
Ṣafā' al-zihn	Item 21	30	1	1.000
	Item 22	30	1	0.900
	Item 23	30	1	0.967
	Item 24	29	0	0.867

Note: Universal Agreement 1 = All participants agree; 0= not all participants agree.

As shown in Table 3, the majority of items demonstrated high levels of agreement among raters, with values well above 0.80. Although the scale-level Universal Agreement (S-FVI/UA) value was 0.54, S-FVI/UA was expected given the stringent 100% agreement requirement across raters on each item. In practice, S-FVI/Ave (0.904) provides a more informative measure of overall clarity. To complement the quantitative evidence from the Face Validity Index (FVI), cognitive interviews were conducted with 15 participants to further assess the clarity and interpretability of the draft items. Respondents generally found the items understandable, though several required one or two rereadings before selecting the most appropriate option.

During the response-process validation stage, most items underwent minor editorial refinements to improve semantic precision, eliminate ambiguous or uncommon terminology, and ensure a clear progression across response options. Two items (Items 3 and 6) were removed due to low I-FVI values and recurrent interpretive difficulties identified in interviews. Following these revisions, the full set of items was submitted to a psychometric expert for finalization, who reviewed the revised wording, content relevance, and alignment with the conceptual definition of ḥikmah. The response processes procedures confirmed that the scale items were precise in representing the

construct of *ḥikmah*, minimized response bias, and ensured readability across respondents, thereby preparing the instrument for subsequent factor analytic procedures (Yusoff, 2019; Gültürk, 2024).

Internal Structure

Preliminary analyses confirmed that the data were suitable for factor analysis. The Kaiser–Meyer–Olkin (KMO) statistic was 0.888, which falls into the "meritorious" category (≥ 0.80) and well above the minimum threshold of 0.70, indicating sampling adequacy. Bartlett's test of sphericity was significant, χ^2 = 1362.930, p < .001, suggesting sufficient inter-item correlations for factor extraction (Lloret et al., 2017; Sigudla & Maritz, 2023). These results supported the continuation of exploratory factor analysis (EFA). The exploratory factor analysis yielded a 19-item structure after removing five items that failed to meet psychometric criteria. Two items (5 and 7) were eliminated due to cross-loading, reflected in factor loadings ≥ .32 on multiple factors, indicating that they did not uniquely represent a single latent construct. Two additional items (15 and 21) had insufficient factor loadings (<.32), suggesting weak associations with their intended factors. Another item (11) was removed because it stood alone on a factor comprising fewer than three items—a configuration considered unstable and psychometrically unreliable.

Table 4

Factor Loading EFA

| Dimension	Items	Factor Loading				M	SD	Item Total Correlation
		1	2	3	4			
Problem Solving	1	0.595				2.2226	0.49905	0.457
	2	0.615				2.2947	0.56719	0.519
	3	0.703				2.2539	0.53327	0.474
	4	0.641				2.3260	0.59371	0.444
	22	0.563				2.4545	0.58031	0.550
	23	0.705				2.4326	0.57270	0.620
	24	0.529				2.5737	0.64919	0.498
Learning	12		0.594			2.4639	0.57005	0.493
	13		0.556			2.4639	0.59170	0.441
	14		0.603			2.5204	0.60817	0.457
	16		0.380			2.4451	0.64620	0.362
	17		0.672			2.3386	0.69008	0.390
Reflecting	18			0.689		2.7806	0.50357	0.429
	19			0.718		2.7367	0.54334	0.498
	20			0.732		2.6144	0.60805	0.447
Understanding	6				0.366	2.6740	0.66846	0.250
	8				0.539	2.5016	0.62382	0.406
	9				0.566	2.5016	0.61876	0.327
	10				0.797	2.7774	0.48628	0.241

Using principal component extraction with promax rotation, the analysis was conducted under the assumption that the underlying factors are conceptually related. Promax rotation is appropriate in this context because it allows the factors to correlate, making it suitable for psychological constructs that are theoretically interconnected (Ahmed & Maruod, 2025). Four factors with eigenvalues greater than 1 were retained. Together, these factors explained 46.96% of the total variance: Factor 1 = 28.33%, Factor 2 = 7.25%, Factor 3 = 5.93%, and Factor 4 = 5.47%. Factor loadings ranged from

0.366 to 0.797, exceeding the recommended minimum of 0.32 (Güvendir & Özkan, 2022). The resulting four-factor solution was labeled *Problem Solving*, *Learning*, *Reflecting*, and *Understanding*, comprising 19 items (see Table 4).

The internal consistency of the ISW-GV was evaluated using McDonald's Omega (ω), computed from the standardized factor loadings. The overall scale demonstrated excellent reliability ($\omega = .919$), while each domain showed acceptable internal consistency: *Problem Solving* ($\omega = .816$), *Learning* ($\omega = .700$), *Reflecting* ($\omega = .756$), and *Understanding* ($\omega = .663$). Although the ω value for the Understanding domain is slightly below the conventional .70 threshold, this does not undermine the psychometric adequacy of the instrument. The ISW-GV is theoretically interrelated facets of the broader construct of ḥikmah rather than standalone subscales. Consequently, modest subdomain reliability coefficients remain acceptable when full-scale reliability is strong (Karabacak Çelik, 2025).

Figure 1.

Model structure of ISW-GV
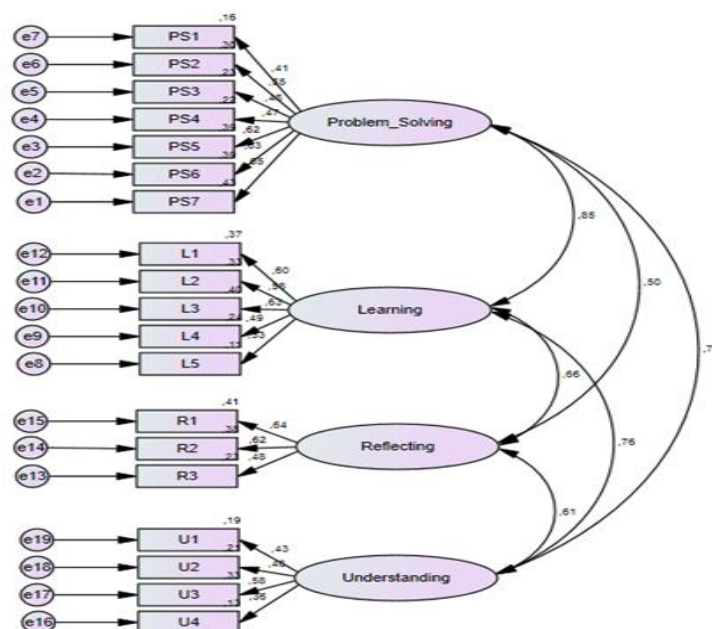
The four-factor model was subsequently tested using confirmatory factor analysis (CFA). Although the chi-square statistic was significant, $\chi^2 = 232.491$, consistent with its sensitivity to sample size, alternative fit indices indicated acceptable model fit: RMSEA = 0.034, GFI = 0.953, TLI = 0.942, and CFI = 0.951 (Hu & Bentler, 1999; Whittaker, 2012; Garnier-Villarreal & Jorgensen, 2025). Factor loadings ranged between 0.330 and 0.654, with no items falling below the 0.30 threshold. The final version of the scale, therefore, consisted of 19 items (see Figure 1).

Inter-factor correlations provided discriminant evidence, as none exceeded the 0.85 cut-off (Cheung et al., 2023). Specifically, correlations were 0.849 between *Problem Solving* and *Learning*, 0.663 between *Learning* and *Reflecting*, 0.612 between *Reflecting* and *Understanding*, 0.496 between *Problem Solving* and *Reflecting*, 0.764 between *Learning* and *Understanding*, and 0.703 between *Problem Solving* and *Understanding*. Taken together, the results of the EFA and CFA confirmed a four-factor internal structure for the Islamic Scale of Wisdom – General Version (ISW-GV), comprising the domains of *Problem Solving*, *Learning*, *Reflecting*, and *Understanding*.

Criterion-Related Evidence

Criterion-related evidence of the ISW-GV was examined by correlating its scores with the Subjective Happiness Scale (SHS) as an external indicator of subjective well-being. Consistent with theoretical expectations, small-to-moderate, positive, significant correlations were considered evidence of criterion validity (Ward et al., 2009).

Table 5

Correlation between ISW-GV, ISW-AV, SHS, GPA, and BSDS among University Students

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. ISW-GV	1														
2. PS	.724**	1													
3. Lr	.699**	.396**	1												
4. Rf	.653**	.199*	.262**	1											
5. Un	.664**	.271**	.233*	.409**	1										
6. ISW-AV	.489**	.486**	.434**	.149	.225*	1									
7.Cm	.463**	.467**	.439**	.139*	.177	.812**	1								
8.Ct	.094	.146	.078	.012	-.001	.518**	.119	1							
9.PS	.456**	.410**	.431**	.112	.267**	.794**	.698**	.088	1						
10.Sch	.340**	.361**	.226**	.077	.242*	.708**	.480**	.242*	.620**	1					
11.Fr	.296**	.399**	.360**	-.007	-.009	.644**	.489**	.230*	.451**	.316**	1				
12.Stb	.365**	.148	.181	.359**	.351**	.545**	.307**	.254**	.435**	.386**	.134	1			
13.SHS	.068	.048	.029	.195*	-0.09	.224*	.133	.196*	.156	.121	.125	.171	1		
14.GPA	-.008	-.005	-.059	-.046	.083	.050	.086	-.172	.120	.163	.048	.042	-.116	1	
15.BSDS	.246*	.071	0.182	.197*	.228*	.321**	.296**	.103	.291**	.252**	.266**	.135**	.194*	.107	1

Note : * $p < .05$, ** $p < .01$; ISW-GV = Islamic Scale of Wisdom – General Version; PS = Problem Solving; Lr = Learning; Rf = Reflecting; Un = Understanding; ISW-AV = Islamic Scale of Wisdom – Academic Version; Cm = Comprehension; Ct = Control, Sch : Scholastic; FR : Fast Reasoning; Stb = Stability; SHS = Subjective Happiness Scale; GPA = Grade Point Average; BSDS= Brief Social Desirability Scale

As shown in Table 5, the total ISW-GV score was not significantly associated with SHS ($r = .068$, $p > .05$) in the student sample. At the subscale level, however, the *Reflecting* dimension demonstrated a small but significant positive correlation with subjective happiness ($r = .195$, $p < .05$). In contrast, *Problem Solving* ($r = .048$, $p > .05$) and *Learning* ($r = .029$, $p > .05$) showed negligible and non-significant associations, while *Understanding* was weakly negative and non-significant ($r = -.090$, $p > .05$). These findings provided partial support for criterion validity, suggesting that reflective aspects of wisdom may be more directly linked to happiness than procedural or cognitive components in student populations.

Table 6

Correlation between ISW-GV, 3D-WS, SHS, and BSDS among Workers

Variables	1	2	3	4	5	6	7	8	9	10	11
1. ISW-GV	1										
2. PS	.801**	1									
3. Lr	.818**	.546**	1								
4. Rf	.652**	.358**	.405**	1							
5. Un	.567**	.212**	.304**	.260**	1						
6. 3D-WS	.438**	.335**	.378**	.287**	.245**	1					
7. Cognitive	.317**	.232**	.242**	.193**	.249*	.764**	1				
8. Reflective	.451**	.339**	.416**	.262**	.255**	.814**	.430**	1			
9. Affective	.204**	.177**	.182**	.191**	.022	.671**	.228**	.383**	1		
10. SHS	.220**	.202**	.178**	.160*	.074	.114	-.045	.172**	.146**	1	
11. BSDS	.044	.025	-.033	.164*	.045	.070	.068	.021	.043	.042	1

Note: * $p < .05$, ** $p < .01$; ISW-GV = Islamic Scale of Wisdom – General Version; PS = Problem Solving; Lr = Learning; Rf = Reflecting; Un = Understanding; 3DWS = Three Dimensional Wisdom Scale; SHS = Subjective Happiness Scale; BSDS = Brief Social Desirability Scale

Results differed in the worker subsample (Table 6). Here, the ISW-GV total score correlated positively and significantly with SHS ($r = .220$, $p < .01$). Dimension-level analyses further supported this association: *Problem Solving* ($r = .202$, $p < .01$), *Learning* ($r = .178$, $p < .01$), and *Reflecting* ($r = .160$, $p < .05$) all correlated significantly with subjective happiness, while *Understanding* remained non-significant ($r = .074$, $p > .05$). These results indicated that in occupational contexts, multiple components of wisdom—not only reflective but also problem-solving and learning contribute to subjective happiness. To more clearly illustrate the contrast in correlation patterns between university students and workers, we present the comparative correlation table between ISW-GV and subjective happiness below.

Table 7

Correlation Patterns of ISW-GV and Happiness Across University Students and Workers

ISW-GV	Correlation with Happiness	
	University Students	Workers
1. Overall	.068	.220**
2. Problem Solving	.048	.202**
3. Learning	.029	.178**
4. Reflecting	.195*	.160*
5. Understanding	-.090	.074

Note: * $p < .05$, ** $p < .01$

Together, these findings suggested that the criterion-related validity of the ISW-GV varies by population. Among students, reflective processing is the primary predictor of happiness, whereas among workers, broader facets of wisdom appear to contribute. Academic environments reward cognitive and reflective depth, while occupational settings require integrated problem solving, learning, and reflective judgment for well-being. Thus, the ISW-GV provides criterion validity evidence that is context-sensitive and consistent with prior research linking wisdom and happiness (Cheng et al., 2025; Siah, 2025).

Convergence Evidence

Convergent evidence was examined by correlating ISW-GV scores with established measures of wisdom, including the *Hikmah Subscale – Noble Character* (HS; Utami, 2020), the Islamic Scale of Wisdom – Academic Version (ISW-AV; Nurtjahjo & Rusdi, 2018), and the Three-Dimensional Wisdom Scale (3D-WS; Ardelt, 2003).

Table 8

Correlation between ISW-GV, 3D-WS, SHS, and BSDS among High School Students

Variables	1	2	3	4	5	6	7	8	9
1. ISW-GV	1								
2. Problem Solving	.778**	1							
3. Learning	.781**	.498**	1						
4. Reflecting	.665**	.325**	.384**	1					
5. Understanding	.680**	.380**	.334**	.305**	1				
6. Hikmah Subscale	.400**	.346**	.302**	.275**	.234**	1			
7. Musyawarah	.422**	.358**	.327**	.291**	.246**	.904**	1		
8. Seimbang	-.115	-.059	-.047	-.105	-.135	-.088	-.110	1	

| 9. BSDS | -.044 | .013 | -.017 | -.114 | -.068 | -.040 | -.020 | .315** | 1 |

Note: * $p < .05$, ** $p < .01$

As shown in Table 8, ISW-GV scores were positively and significantly correlated with the Hikmah Subscale ($r = .400$, $p < .01$). At the dimension level, *Problem Solving* ($r = .346$, $p < .01$), *Learning* (r = .302, p < .01), *Reflecting* (r = .275, p < .01), and *Understanding* ($r = .234$, $p < .01$) were all positively associated with the Hikmah Subscale. These findings confirmed strong convergent evidence in this population, as both the total and subscale scores of ISW-GV aligned with an external measure designed specifically for high school students. Notably, the *Musyawarah* (deliberative) subdimension of the Hikmah Subscale showed the strongest correlation, whereas the *Seimbang* (balance) subdimension did not demonstrate significant associations, suggesting differences in operationalization at the behavioral level.

Additional convergent evidence was obtained by correlating ISW-GV with the ISW-AV. As reported in Table 5, the ISW-GV total score correlated moderately with ISW-AV ($r = .489$, $p < .01$). Dimension-level correlations revealed significant associations for *Problem Solving* ($r = .486$, $p < .01$), *Learning* ($r = .434$, $p < .01$), and *Understanding* ($r = .225$, $p < .05$), while *Reflecting* showed a weak and non-significant association ($r = .149$, $p > .05$). These results indicated a stronger similarity between the two instruments on cognitive and procedural aspects of wisdom, whereas reflective processes may be operationalized differently across the scales.

Convergent evidence was further assessed against the 3D-WS in the worker sample (Table 6). ISW-GV total scores were moderately correlated with 3D-WS ($r = .438$, $p < .01$). Subscale correlations were also significant: *Problem Solving* ($r = .335$, $p < .01$), *Learning* ($r = .378$, $p < .01$), *Reflecting* ($r = .287$, $p < .01$), and *Understanding* ($r = .245$, $p < .01$). The magnitude of these associations, ranging from small to moderate, is consistent with theoretical expectations for related but non-identical constructs. Stronger correlations in cognitive and learning aspects suggest partial overlap between ISW-GV and the 3D-WS, while reflective and affective dimensions appear to be uniquely conceptualized within the Islamic framework of *ḥikmah*. Taken together, these findings provide robust evidence of convergent evidence across different populations. The ISW-GV demonstrates meaningful associations with existing wisdom measures while maintaining distinctive features rooted in Islamic ethical philosophy.

Discriminant Evidence

Discriminant evidence was examined by correlating ISW-GV scores with Grade Point Average (GPA), a measure of academic achievement. Theoretically, wisdom (*ḥikmah*) emphasizes proportional reasoning, moral reflection, and practical judgment, whereas GPA primarily reflects factual and analytical knowledge. Accordingly, low correlations were expected, as wisdom and academic achievement represent distinct constructs (Fiori & Antonakis, 2011). As shown in Table 5, correlations between ISW-GV and GPA were uniformly small and nonsignificant. The total ISW-GV score correlated at $r = -.008$, while subscale correlations were $r = -.005$ for *Problem Solving*, $r = -.059$ for *Learning*, $r = -.046$ for *Reflecting*, and $r = .083$ for *Understanding*. None of the coefficients approached a medium effect size. This pattern provides strong evidence of discriminant validity. ISW-GV does not reduce to academic achievement; rather, it captures distinct aspects of proportional reasoning, reflective judgment, and moral discernment that GPA does not explain.

Examining the Impact of Social Desirability Bias

Assessing the influence of social desirability is a crucial aspect of instrument validation, as it helps determine whether responses are distorted by the tendency to provide socially desirable answers (Lanz et al., 2022). Correlations of approximately .10, .20, and .30 are interpreted as small, moderate, and large, respectively (Funder & Ozer, 2019). In this study, correlations below .30 were taken as evidence that the ISW-GV was minimally affected by social desirability bias.

Table 9

Correlation Patterns of ISW-GV and Social Desirability Across High School Students, University Students, and Workers

ISW-GV	Correlation with Social Desirability		
	High School Students	University Students	Workers
1. Overall	-.044	.246*	.044
2. Problem Solving	.013	.071	.025
3. Learning	-.017	.182	-.033
4. Reflecting	-.114	.197*	.164*
5. Understanding	-.068	.228*	.045

Note: * $p < .05$, ** $p < .01$

In the university student sample, as reported in Table 9, correlations between the ISW-GV and the Brief Social Desirability Scale (BSDS) were small. The total ISW-GV score correlated at $r = .246$ with BSDS, while dimension-level correlations were $r = .071$ for *Problem Solving*, $r = .182$ for *Learning*, $r = .197$ for *Reflecting*, and $r = .228$ for *Understanding*. All coefficients remained below the .30 threshold, indicating a low impact of social desirability bias. In the high school sample (Table 9), all correlations were near zero and nonsignificant. The ISW-GV total score correlated negatively with BSDS ($r = -.044$), with dimension-level correlations ranging from $r = -.114$ (*Reflecting*) to $r = .013$ (*Problem Solving*). These findings provide strong evidence of discriminant validity, confirming that the ISW-GV does not simply capture "socially desirable" responses in adolescent populations.

Among workers (Table 9), correlations between ISW-GV and BSDS also remained small. The total score correlated at $r = .044$, with subscale correlations ranging from $r = -.033$ (*Learning*) to $r = .164$ (*Reflecting*). The observed range ($r = -.033$ to .164) further indicates minimal susceptibility to social desirability effects in occupational contexts. Across all three samples, correlations between ISW-GV and BSDS were consistently small ($r = -.044$ to .246), well below the .30 threshold. These results suggested that the ISW-GV is not substantially influenced by socially desirable responding, thereby strengthening its construct validity.

Integrating Ibn Miskawaih's Seven Derivatives of *Ḥikmah* into Four Empirical Factors

The factor analysis of the Islamic Scale of Wisdom – General Version (ISW-GV) revealed a coherent four-factor structure comprising *Problem Solving, Learning, Reflecting,* and *Understanding*. Although Ibn Miskawaih conceptualized seven cognitive–reflective qualities within the broader virtue of *ḥikmah*, empirical consolidation of these attributes into four factors demonstrates how related theoretical components cluster together in actual response patterns. This integration is consistent with contemporary psychometric perspectives, which show that conceptually adjacent traits often converge into broader functional domains when empirically assessed in modern populations (Glück & Weststrate, 2022). Accordingly, each factor represents an empirically distinguishable yet theoretically grounded manifestation of *ḥikmah*, retaining the substantive essence of Miskawaih's framework while enhancing measurement parsimony and structural stability. The

following subsections elaborate on how each empirical factor reflects the underlying theoretical elements of ḥikmah and how these domains align with current wisdom research.

The first factor, *Problem Solving*, is defined as the capacity to derive appropriate solutions and articulate them clearly and accessibly. Empirically, this factor integrates items originally derived from *al-dzakā'* and *ṣafā' al-zihn*. Items loading on this factor capture both the speed and accuracy with which individuals draw inferences, consider consequences, and formulate resolutions, as well as their ability to communicate these conclusions in a structured, understandable manner. This structure remains faithful to Ibn Miskawaih's view that *al-dzakā'* is not merely rapid inference but also involves grasping causal relations and generating effective solutions, while *ṣafā' al-zihn* reflects mental clarity that allows these solutions to be expressed in a calm, coherent form. Contemporary wisdom research likewise highlights problem-focused reasoning and the formulation of context-sensitive solutions as central components of wise functioning in everyday life and complex social contexts (Santos & Grossmann, 2021).

The second factor, *Learning*, is defined as the ability to efficiently grasp new information and retain what has been learned. This factor combines items derived from *surʿah al-fahm* and *al-dhikr*, focusing on how individuals process, integrate, and later recall theoretical knowledge, situations, and problems. In Ibn Miskawaih's framework, *surʿah al-fahm* refers to rapid and accurate comprehension, whereas *al-dhikr* involves the stable retention and retrieval of beneficial knowledge and experiences. Collapsing these into a single learning factor is conceptually coherent because, in a psychological sense, *Learning* presupposes both initial understanding and subsequent recall (Zhang et al., 2023). Recent empirical work on wisdom emphasizes that lifelong learning, cognitive flexibility, and the capacity to integrate new information with prior knowledge are key characteristics of wise individuals, particularly in dynamic and uncertain environments (Mascolo & Stammberger, 2024).

The third factor, *Reflecting*, is defined as the capacity to deeply contemplate past experiences and derive meaningful lessons from them. This factor corresponds directly to *jaudah al-zihn*, which Ibn Miskawaih describes as the soul's ability to review and evaluate previous actions and events in a balanced way. The items that form this factor focus on how individuals engage in self-examination, evaluate their past decisions, and transform experiences into enduring insights. This is consistent with contemporary models that treat reflective or metacognitive processing as a core component of wisdom, enabling individuals to reconsider their assumptions, recognize their limitations, and refine their judgments over time (Stichter, 2024).

The fourth factor, *Understanding*, is defined as the ability to filter information and form objective judgments about its truth and value. This factor integrates items derived from *al-taʿaqqul* and *suhūlah al-taʿallum*. In Ibn Miskawaih's account, *al-taʿaqqul* denotes thinking that is grounded in appropriate evidence and correspondence with reality, while *suhūlah al-taʿallum* reflects an ease in comprehending even abstract or conceptual material, coupled with discernment regarding its quality. In our empirical model, these two derivatives converge into a coherent factor that emphasizes fair-minded evaluation, critical analysis, and openness to revising one's views in the face of better evidence. This aligns with contemporary wisdom literature, which frames wise understanding as a blend of epistemic humility, critical thinking, and context-sensitive judgment rather than the sheer accumulation of knowledge (Sternberg et al., 2021). Together, these four empirically derived factors preserve the substantive content of Ibn Miskawaih's seven derivatives of ḥikmah while organizing them into a parsimonious structure that is psychometrically robust and theoretically interpretable.

This structure aligns closely with Jakubik's (2025) multidimensional framework of wisdom, which includes knowing ourselves, connecting to others, knowing the world, and connecting to the

world. Specifically, *Reflecting* parallels with knowing ourselves, emphasizing introspection and learning from experience; *Learning* corresponds to knowing the world through openness to new knowledge; *Problem-Solving* is consistent with connecting to the world by addressing real-life decision-making and problem resolution; and *Understanding* overlaps with both knowing the world and connecting to others, highlighting contextual comprehension and clear communication. This convergence demonstrates that the ISW-GV dimensions are not only rooted in the Islamic tradition of ḥikmah but also resonate with broader, contemporary theories of wisdom, thereby strengthening both the conceptual validity and the cross-context relevance of the scale.

Convergence of Islamic and Western Wisdom Measures

The moderate correlation between ISW-GV and ISW-AV suggests that wisdom is a consistent construct across both academic and everyday contexts (Glück & Weststrate, 2022). While the two scales share a common conceptual core, each retains contextual specificity. The ISW-AV emphasizes wisdom expressed in academic performance, whereas the ISW-GV captures manifestations of wisdom in broader life experiences. This distinction reflects the appropriateness of contextualizing wisdom assessment without compromising the integrity of the underlying construct (Grossmann et al., 2020).

Comparisons with established Western measures also yielded meaningful insights. The moderate correlations between ISW-GV and the Three-Dimensional Wisdom Scale (3D-WS; Ardelt, 2003) suggest shared constructs while preserving the theoretical distinctiveness. ISW-GV's *Problem Solving* dimension is conceptually related to the reflective domain of the 3D-WS, emphasizing evaluation of alternatives before decision-making. Similarly, ISW-GV's *Learning* and *Understanding* align with the cognitive dimension, reflecting knowledge acquisition, comprehension, and integration of information. The Reflecting factor in ISW-GV also aligns with the reflective domain of the 3D-WS, which emphasizes introspection and meaning-making. ISW-GV highlights cognitive-rational processes within the framework of Islamic ethics, while the 3D-WS explicitly incorporates affective aspects. This pattern supports the view that wisdom has universal core features while remaining adaptable across diverse cultural and value-based contexts (Lin et al., 2025).

Life Experiences Moderate Wisdom and Happiness

Among university students, ISW-GV scores did not correlate significantly with subjective happiness, suggesting that the relationship between wisdom and well-being may not be linear, particularly in younger populations where wisdom is still developing and not fully integrated with lived experience (Puchalska-Wasyl, 2023). By contrast, among workers, ISW-GV scores were positively associated with happiness across total and subscale scores. This pattern is consistent with longitudinal research indicating that wisdom is a stronger predictor of subjective well-being in adulthood as individuals accumulate complex life experiences and greater responsibilities that necessitate reflective judgment and self-regulation (Santos & Grossmann, 2021; Zhang et al., 2024).

Contemporary lifespan research shows that wisdom-related judgment becomes more consequential for happiness as individuals accumulate emotionally complex, morally challenging, and socially significant experiences that require long-term reflection and adaptive regulation (Glück & Weststrate, 2022). Such experiences provide the "raw material" through which cognitive, reflective, and moral components of wisdom can be enacted in daily life, thereby allowing wisdom to exert measurable benefits on well-being. Younger individuals, such as university students, may possess early forms of wisdom-related capacities, but these capacities often remain latent or weakly

expressed because they have not yet been tested against high-stakes life events or sustained responsibilities. In contrast, adults with richer life histories can apply their wisdom more consistently to manage stress, resolve interpersonal tensions, and derive meaning from adversity, resulting in stronger associations with subjective happiness (Cheng et al., 2025).

From Bipolar Format to Rank-Order Response Format

Despite these strengths, several methodological limitations warrant consideration. The bipolar item format of the ISW-GV, which situates responses between two extremes such as foolishness versus rashness, may introduce psychometric ambiguity. It is not always clear whether respondents are rejecting one pole, endorsing the other, or positioning themselves relative to both. Future studies should therefore consider disentangling wisdom, rashness (*al-safh/khid'ah*), and foolishness (*balāhah*) into separate constructs, each assessed independently. Such an approach would allow for more precise measurement of these distinct yet interrelated moral dispositions. A promising methodological refinement involves adapting the instrument into a Situational Judgment Test (SJT) format with rank-order responses (Whetzel et al., 2020; Wolcott et al., 2021). Instead of selecting from options with dual anchors, participants could rank potential responses from the most to the least wise. This design would enhance ecological validity by requiring respondents to evaluate alternatives in real-life scenarios, reduce the risk of socially desirable responding, and capture individual tendencies across the three value poles of wisdom (*ḥikmah*), rashness (*sufh*), and foolishness (*balāh*). Implementing such a format would not only address psychometric concerns but also deepen the scale's conceptual clarity and cultural resonance.

CONCLUSION

The present study established the Islamic Scale of Wisdom - General Version (ISW-GV) as a psychometrically robust instrument grounded in classical Islamic virtue ethics while demonstrating strong empirical validity across diverse samples. The scale demonstrated strong content validity grounded in Ibn Miskawaih's conceptual framework of wisdom and produced a stable four-factor structure (*Problem-Solving, Learning, Reflecting*, and *Understanding*), as confirmed by CFA. Reliability indices were acceptable. Convergent validity was supported through moderate correlations with both ISW-AV, HS, and 3D-WS, indicating conceptual alignment with previously established measures of wisdom across Islamic and Western frameworks. Criterion-related evidence showed differentiated patterns: wisdom was only weakly linked to subjective happiness among students but demonstrated small-medium yet consistent associations among workers, suggesting that wisdom becomes more consequential for well-being as life experience deepens. Discriminant validity was confirmed through near-zero correlations with academic achievement (GPA) and consistently low associations with social desirability, indicating that ISW-GV captures morally reflective reasoning rather than socially desirable responding or merely academic ability.

Overall, these results supported ISW-GV as a valid and culturally relevant instrument for assessing wisdom in the general population. However, the current response format may limit the ability to clearly distinguish between tendencies toward wisdom, rashness, and stupidity. Future versions of the instrument would benefit from separating these constructs into distinct but related dimensions, preferably using a rank-order Situational Judgment Test (SJT) format. This would enable respondents not only to select but also to rank possible actions from most to least wise, providing deeper insight into their moral decision-making patterns. With these improvements, ISW-GV could evolve beyond a traditional rating scale into a more realistic tool for understanding ethical judgment

across different value orientations. Future research should also validate the ISW-GV across other cultural settings and Muslim-majority countries to determine whether the Islamic conceptualization of wisdom generalizes beyond the Indonesian context.

AUTHORS' CONTRIBUTIONS: AR designed the study, supervised the data analysis, and prepared the final manuscript. FEN contributed to the literature review and co-writing of the final manuscript. NA examined the conceptualization of ḥikmah within Islamic thought and conducted interviews with Islamic scholars. IM provided evidence for content validity. RKP contributed evidence for response process validity. AS performed the exploratory and confirmatory factor analyses. AVPA gathered evidence for criterion-related validity. All authors have read and approved the final version of the manuscript.

CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

REFERENCES

Ahmed, A., & Maruod, W. (2025). Comparative analysis of varimax and Promax rotation methods in exploratory factor analysis. *Edelweiss Applied Science and Technology*, *9*(5), 501–513. https://doi.org/10.55214/25768484.v9i5.6929

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Exploratory factor analysis and principal component analysis in clinical studies: Which one should you use? *Journal of Advanced Nursing*, *76*(8). https://doi.org/10.1111/jan.14377

Alfaiz, A., Reza, I. F., Elsyad, M. K., Inayah, A. N., Apriliyanti, P., Putri, A. A., Anggraeni, T., Rafiola, R. H., Thahir, M., & Da Costa, A. (2025). The role of religious wisdom in shaping academic culture among Indonesian university students: A descriptive study with MANOVA analysis. *Psikis: Jurnal Psikologi Islami*, *11*(1), 126–136. https://doi.org/10.19109/psikis.v11i1.27814

Ardelt, M. (2003). Empirical assessment of a Three-Dimensional Wisdom Scale. *Research on Aging*, *25*(3), 275–324. https://doi.org/10.1177/0164027503251764

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*(1), 79–86. https://doi.org/10.1177/0748175613513808

Bergsma, A., & Ardelt, M. (2012). Self-reported wisdom and happiness: An empirical investigation. *Journal of Happiness Studies*, *13*(3), 481–499. https://doi.org/10.1007/s10902-011-9275-5

Cheng, S.-L., Zhang, X., Liu, S., Zhao, C., & Bi, S. (2025). Integration and completion: Life wisdom and happiness of Chinese older adults from a life course perspective. *Frontiers in Public Health*, *13*, 1575099. https://doi.org/10.3389/fpubh.2025.1575099

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2023). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*(2), 745–783. https://doi.org/10.1007/s10490-023-09871-y

Fiori, M., & Antonakis, J. (2011). The ability model of emotional intelligence: Searching for valid measures. *Personality and Individual Differences*, *50*(3), 329–334. https://doi.org/https://doi.org/10.1016/j.paid.2010.10.010

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156 – 168. https://doi.org/10.1177/2515245919847202

Garnier-Villarreal, M., & and Jorgensen, T. D. (2025). Evaluating local model misspecification with modification indices in bayesian structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *32*(2), 304–318. https://doi.org/10.1080/10705511.2024.2413128

Glück, J., & Weststrate, N. M. (2022). The wisdom researchers and the elephant: An integrative model of wise behavior. *Personality and Social Psychology Review*, *26*(4), 342–374. https://doi.org/10.1177/10888683221094650

Grossmann, I., Weststrate, N. M., Ardelt, M., Brienza, J. P., Dong, M., Ferrari, M., Fournier, M. A., Hu, C. S., Nusbaum, H. C., & Vervaeke, J. (2020). The Science of wisdom in a polarized world: Knowns and unknowns. *Psychological Inquiry*, *31*(2), 103–133. https://doi.org/10.1080/1047840X.2020.1750917

Gültürk, E. A. (2024). Scale adaptation and redevelopment: A review on validity and reliability. *Journal of Cellular and Molecular Immunology*, *3*(1), 26–32. https://doi.org/10.46439/immunol.3.027

Güvendir, M. A., & Özkan, Y. Ö. (2022). Item removal strategies conducted in exploratory factor analysis: A comparative study. *International Journal of Assessment Tools in Education*, *9*(1). https://doi.org/10.21449/ijate.827950

Haghighat, R. (2007). The development of the Brief Social Desirability Scale (BSDS). *Europes Jpurnal of Psychology*, *3*(4). https://doi.org/https://doi.org/10.5964/ejop.v3i4.417

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ibn Miskawaiḥ. (2011). *Tahdhīb al-akhlāq (The refinement of character)* (al-Halālī, Ed.). Manshūrāt al-Jamal. http://archive.org/download/745789549844/tahtheb-alakhlaq.pdf

Indati, A., Adiyanti, M. G., & Ramdhani, N. (2019). Peran wisdom terhadap life satisfaction pada lansia. *Gadjah Mada Journal of Psychology (GamaJoP)*, *5*(1), 60–69. https://doi.org/10.22146/gamajop.47967

Jakubik, M. (2025). Rethinking the dimensions of wisdom in higher educational pedagogy: Grounded theory research. *Trends in Higher Education*, *4*(2). https://doi.org/10.3390/higheredu4020020

Jifa, G. (2013). Data, information, knowledge, wisdom and meta-synthesis of wisdom-comment on wisdom global and wisdom cities. *Procedia Computer Science*, *17*, 713 – 719. https://doi.org/10.1016/j.procs.2013.05.092

Karabacak Çelik, A. (2025). A reliability generalization meta-analysis of the PERMA-profiler. *Social Indicators Research*, *177*(3), 1027–1050. https://doi.org/10.1007/s11205-025-03547-4

Kaymak, S., Basaran-Acil, S., & Duygulu, S. (2025). Psychometric properties of the self-assessed wisdom scale among Turkish senior nursing students. *BMC Nursing*, *24*(1). https://doi.org/10.1186/s12912-025-03355-8

Khan, M. I. (2025). Wisdom and virtue ethics in the Cairo Genizah: The intersection of Jewish and Islamic moral education. *Entangled Religions*, *15*(3). https://doi.org/10.46586/er.15.2025.11889

Krieger, F., & Greiff, S. (2023). Intelligence, creativity, and wisdom: A case for complex problem solving? In *Intelligence, Creativity, and Wisdom: Exploring their Connections and Distinctions* (pp. 225 – 250). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-26772-7_10

Lanz, L., Thielmann, I., & Gerpott, F. H. (2022). Are social desirability scales desirable? A meta-analytic test of the validity of social desirability scales in the context of prosocial behavior. *Journal of Personality*, *90*(2). https://doi.org/10.1111/jopy.12662

Lin, H., Fu, X., & Wang, J. F. (2025). Reliability and validity of the Integrated Wisdom Scale (long- and short-form) for adolescents in a Chinese sample. *Journal of Research on Adolescence*, *35*(3). https://doi.org/10.1111/jora.70060

Liu, Y., Lyu, D., Xie, S., Yao, Y., Liu, J., Lu, B., Zhang, W., Xian, S., Yan, J., Gong, M., Wu, X., Li, Y., Zhang, H., Zhou, J., Zhou, Y., Lin, M., Yin, H., Wang, X., Wang, Y., … Pan, X. (2025). The impact of grade point average on medical students' perception of the learning environment: a

multicenter cross-sectional study across 12 Chinese medical schools. *BMC Medical Education*, *25*(1). https://doi.org/10.1186/s12909-025-06977-3

Lloret, S., Ferreres, A., Hernández, A., & Tomás, I. (2017). The exploratory factor analysis of items: Guided analysis based on empirical data and software. *Anales de Psicologia*, *33*(2), 417–432. https://doi.org/10.6018/analesps.33.2.270211

Lyubomirsky, S., & Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, *46*(2), 137–155. https://doi.org/10.1023/A:1006824100041

Mahpur, M., Rosdiana, A. M., Wafa, I. A., Jamilah, J., & Pramitha, D. (2023). Perception and role of lecturers' practical wisdom in the initial implementation of religious moderation for general courses. *Psikis: Jurnal Psikologi Islami*, *9*(2), 351–367. https://doi.org/10.19109/psikis.v9i2.18767

Mascolo, M. F., & Stammberger, I. (2024). Bridging the fact/value divide in wisdom research: The development of expertise in wise decision-making. *Topoi*, *43*(3), 937–949. https://doi.org/10.1007/s11245-023-10006-7

Nabi, U. (2024). Akhlāq and subjective well-being: Exploring the mediating role of perceived stress using structural equation modelling. *Social Sciences and Humanities Open*, *10*. https://doi.org/10.1016/j.ssaho.2024.101000

Nurtjahjo, F. E., & Rusdi, A. (2018). Developing the Islamic Scale of Wisdom: Academic version (ISW-AV). *Middle East Journal of Positive Psychology*, *4*, 12–24. https://www.middleeastjournalofpositivepsychology.org/index.php/mejpp/article/view/62

Nurtjahjo, F. E., Rusdi, A., Widi Anggraeni, M., Hanifah, I., Nur Ummi, F., Azizah Destyaputri, H., & Zulfa Hardiana, A. (2021). Validasi Islamic Scale of Wisdom-Academic Version dengan San Diego Wisdom Scale dan Papi-Kostick. *Psikologika: Jurnal Pemikiran Dan Penelitian Psikologi*, *26*(2). https://doi.org/10.20885/psikologika.vol26.iss2.art10

Omar, M. N., Long, A. S., Mahmood, A. R., & Rahman, Z. A. B. (2021). Islamic notion of happiness (al-Sa'ada): An analysis of Miskawayh's thought. *International Journal of Islamic Thought*, *19*(1), 49 – 57. https://doi.org/10.24035/IJIT.19.2021.195

Othman, M. K. H., Yaakob, M. F. M., Zulhilmi Haron, Mohd., Abdullah, S., Zakaria, N., & Mahzan, S. K. (2021). Exploratory factor analysis and reliability of 'student outcome' instrument using Ibnu Maskawaih's theory. *Central Asia and the Caucasus*, *22*(5), 711 – 728. https://doi.org/10.37178/ca-c.21.5.058

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford University Press.

Puchalska-Wasyl, M. M. (2023). Do wisdom and well-being always go hand in hand? The role of dialogues with oneself. *Journal of Happiness Studies*, *24*(3), 1059–1074. https://doi.org/10.1007/s10902-023-00621-x

Raharja, B. N., & Indati, A. (2019). Kebijaksanaan dan kepuasan hidup pada remaja. *Gadjah Mada Journal of Psychology (GamaJoP)*, *4*(2), 96–104. https://doi.org/10.22146/gamajop.46354

Rusdi, A., Nurasha, K. P., Rachmawati, M. A., Yang, L., Wangnan, P., & Andriansyah, Y. (2022). Models of helping, happiness, and spiritual serenity among Indonesian and Chinese students. *International Journal of Innovative Research and Scientific Studies*, *5*(2). https://doi.org/10.53894/ijirss.v5i2.393

Santos, H. C., & Grossmann, I. (2021). Cross-temporal exploration of the relationship between wisdom-related cognitive broadening and subjective well-being: Evidence from a cross-validated national longitudinal study. *Social Psychological and Personality Science*, *12*(4), 506–516. https://doi.org/10.1177/1948550620921619

Shi, J., Mo, X., & Sun, Z. (2012). Content validity index in scale development. *Journal of Central South University (Medical Sciences)*, *37*(2), 152–2. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007

Siah, P. C. (2025). Non-attachment, sense of coherence and happiness: An examination of their relationships through the Buddhist concept of virtue-meditation-wisdom. *Journal of Spirituality in Mental Health*, *27*(2), 260 – 275. https://doi.org/10.1080/19349637.2024.2342975

Sigudla, J., & Maritz, J. E. (2023). Exploratory factor analysis of constructs used for investigating research uptake for public healthcare practice and policy in a resource-limited setting, South Africa. *BMC Health Services Research*, *23*(1), 1423. https://doi.org/10.1186/s12913-023-10165-8

Sternberg, R. J., Glaveanu, V., Karami, S., Kaufman, J. C., Phillipson, S. N., & Preiss, D. D. (2021). Meta-intelligence: Understanding, control, and interactivity between creative, analytical, practical, and wisdom-based approaches in problem solving. *Journal of Intelligence*, *9*(2). https://doi.org/10.3390/jintelligence9020019

Stichter, M. (2024). Flourishing goals, metacognitive skills, and the virtue of wisdom. *Topoi*, *43*(3), 963–981. https://doi.org/10.1007/s11245-024-10013-2

Utami, E. S. (2020). *Pengembangan dan validasi instrumen akhlak mulia siswa dengan pendekatan item response theory Model RASCH*. Universitas Islam Indonesia.

Vaisi, S., Kordnoghabi, R., Imani, S., & Kashefi, F. (2025). Psychometric properties of the Persian version abbreviated Wisdom Scale in Iranian adults. *Journal of Applied Psychological Research*, *16*(3), 161 – 179. https://doi.org/10.22059/japr.2024.361950.644670

Ward, C., Fischer, R., Lam, F. S. Z., & Hall, L. (2009). The convergent, discriminant, and incremental validity of scores on a self-report measure of cultural intelligence. *Educational and Psychological Measurement*, *69*(1), 85–105. https://doi.org/10.1177/0013164408322001

Whetzel, D., Sullivan, T., & McCloy, R. (2020). Situational judgment tests: An overview of development practices and psychometric characteristics. *Personnel Assessment and Decisions*, *6*(1), 1–16. https://doi.org/10.25035/pad.2020.01.001

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *Journal of Experimental Education*, *80*(1), 26–44. https://doi.org/10.1080/00220973.2010.531299

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, *45*(3), 197–210. https://doi.org/10.1177/0748175612440286

Wolcott, M. D., Lobczowski, N. G., Zeeman, J. M., & McLaughlin, J. E. (2021). Exploring the role of item scenario features on situational judgment test response selections. *American Journal of Pharmaceutical Education*, *85*(6). https://doi.org/10.5688/ajpe8546

Yusoff, M. S. B. (2019). ABC of response process validation and face validity index calculation. *Education in Medicine Journal*, *11*(3), 55–61. https://doi.org/10.21315/eimj2019.11.3.6

Zhang, H., Glassman, H., Ji, L.-J., Huang, C., & Hu, C. S. (2024). Future-oriented temporal perspective promotes wise reasoning. *Journal of Personality*, *92*(4), 1158 – 1171. https://doi.org/10.1111/jopy.12879

Zhang, K., Shi, J., Wang, F., & Ferrari, M. (2023). Wisdom: Meaning, structure, types, arguments, and future concerns. *Current Psychology*, *42*(18), 15030–15051. https://doi.org/10.1007/s12144-022-02816-6