

Beyond Classical Testing: A Hybrid CFA and Rasch Validation of Comprehensive Intellectual Humility Scale in Indonesia

Syahid Izharuddin Lubis

Faculty of Psychology, Universitas Negeri Jakarta, Jakarta, Indonesia

Tina Deviana

Department of Psychology, Faculty of Philosophy and Humanity, Universitas Paramadina, Jakarta, Indonesia

Abstract. Intellectual humility is an important construct for enhancing open-mindedness, but the measurement lacks cross-cultural precision. Therefore, this study aimed to provide a rigorous psychometric evaluation of Comprehensive Intellectual Humility Scale (CIHS) in Indonesian context (N=418), using a dual-framework approach of Confirmatory Factor Analysis (CFA) and Rasch Modelling. The results showed that CFA confirmed a robust four-factor structure with high internal consistency across all dimensions. Furthermore, Rasch analysis validated the scale unidimensionality and person-item reliability. Wright Map analysis reported a consistency gap, where several items were easily endorsed, suggesting a potential ceiling effect in Indonesian population. CIHS was established as a high-precision tool for local study while reporting specific areas for item recalibration to enhance measurement sensitivity through the integration of Classical Test and Item Response Theories.

Keywords: Confirmatory factor analysis, comprehensive intellectual humility scale, intellectual humility, psychometric properties, Rasch model

Presisi Pengukuran Kerendahan Hati Intelektual: Validasi Komprehensif CIHS Versi Indonesia Melalui Pendekatan Teori Tes Klasik dan Modern

Abstrak. Meskipun kerendahan hati intelektual (*intellectual humility*) krusial bagi keterbukaan pikiran dalam masyarakat yang terpolarisasi, ketersediaan instrumen pengukuran yang tervalidasi secara psikometris di Indonesia masih sangat terbatas. Penelitian ini bertujuan untuk menguji validitas dan reliabilitas Comprehensive Intellectual Humility Scale (CIHS) menggunakan pendekatan hibrida: Analisis Faktor Konfirmatori (CFA) dan Model Rasch. Melibatkan 418 partisipan, hasil CFA mengkonfirmasi struktur empat faktor yang fit secara statistik, mendukung konsistensi internal instrumen. Lebih lanjut, analisis Rasch membuktikan bahwa CIHS memenuhi asumsi unidimensionalitas dengan reliabilitas butir dan person yang sangat kuat. Namun, analisis Peta Wright mengungkap adanya ceiling effect pada beberapa butir yang terlalu mudah disetujui oleh partisipan Indonesia, mengindikasikan perlunya kalibrasi tingkat kesulitan butir di masa depan. Temuan ini menegaskan bahwa CIHS adalah instrumen yang tangguh untuk penggunaan klinis dan riset di Indonesia, sekaligus memberikan peta jalan untuk optimasi alat ukur psikologis yang peka terhadap kemampuan responden.

Kata Kunci: Confirmatory factor analysis, comprehensive intellectual humility scale, intellectual humility, psychometric properties, Rasch model

Correspondence: Syahid Izharuddin Lubis. Email: syahid.izharuddin@unj.ac.id

Introduction

Intellectual humility, defined as a non-threatening awareness of fallibility (Krumrei-Mancuso & Rouse, 2016), is a virtue regarded as normatively important (Ballantyne, 2023). In psychology, this awareness is classified as an epistemic virtue and has received increasing scholarly attention in recent years (Alfano et al, 2017; Chancellor & Lyubomirsky, 2013; Hook et al, 2017; Hoyle et al, 2016; Krumrei-Mancuso et al, 2020; Porter, Elnakouri, et al, 2022; Tangney, 2000). The growing interest has led to the development of diverse theoretical frameworks and measurement instruments designed to conceptualize and assess intellectual humility (Alfano et al, 2017; Porter & Schumann, 2018; Zmigrod et al, 2019).

Studies on intellectual humility remain at relatively early stages, with multiple definitions reporting different facets of the multifaceted construct (Hill et al, 2021). Intellectual humility is conceptually distinct from the general context of humility (Davis et al, 2016). Even though general humility includes “an accurate view of strengths and weaknesses” across situations, intellectual humility specifically comprises openness to alternative viewpoints and the recognition of knowledge or beliefs (Krumrei-Mancuso & Rouse, 2016; Leary et al, 2017).

Conceptualizations of intellectual humility can be broadly divided into implicit and explicit theories (Porter, Baldwin, et al, 2022). Implicit conceptions characterize

intellectually humble individuals as possessing cognitive qualities (intelligence, curiosity), self-oriented traits (modesty), and other-oriented characteristics such as politeness and reliability (Barrett, 2017; McElroy et al, 2014). Some incorrectly associate humility with self-deception, self-abasement, low self-esteem, humiliation, or ignorance (Davis et al, 2016; Tangney, 2000). In contrast, explicit theories developed by philosophers and social scientists offer more precise and theoretically grounded accounts of intellectual humility (Krumrei-Mancuso et al, 2020).

Several studies conceptualize intellectual humility as a trait and a virtue. However, some scholars view the concept as a situational state including phenomenological experience (Porter et al, 2022). Across existing definitions, two key dimensions of intellectual humility are consistently identified, namely the intrapersonal (self-focused) and the interpersonal (other-focused). The intrapersonal dimension includes an accurate perception of oneself as a knower, minimal concern for intellectual self-importance or dominance, a non-defensive awareness of intellectual fallibility, a willingness to acknowledge the limitations of beliefs, knowledge, and experiences, and recognition of intellectual strengths and capabilities (Krumrei Mancuso, 2017; Porter et al, 2022; Porter & Schumann, 2018). In contrast, interpersonal dimension (other-focused) comprises appreciation of others’ knowledge and intellectual potential, respect for alternative

viewpoints, and the willingness to engage in interactions with others (Hook et al, 2017; Van Tongeren et al, 2019).

In recent years, empirical studies on intellectual humility have grown significantly, leading to various measurement scales (Haggard et al, 2018; Krumrei-Mancuso et al, 2020). Several examples of the measures are presented in Table 1. This includes Comprehensive Intellectual Humility Scale (CIHS) by Krumrei-Mancuso and

Rouse (Krumrei-Mancuso & Rouse, 2016; Leary et al, 2017). The rationale for selecting the questionnaire is because of the multifaceted framework, which facilitates a comprehensive examination of intellectual humility (Alfano et al, 2017; Zachry et al, 2018). CIHS is among the first self-report instruments designed to measure intellectual humility across intrapersonal and interpersonal domains (Haggard et al, 2018; Krumrei-Mancuso & Rouse, 2016).

Table 1

Measures of Intellectual Humility

Scale Name	Structure	Reliability	Article Sources
Limitations-Owning Intellectual Humility Scale	12 items; 3 subscales	Overall: $\alpha = .86$ Love of learning: $\alpha = .81$ Appropriate discomfort with limitations: $\alpha = .84$ Owning intellectual limitations: $\alpha = .77$	Haggard et al.
Multi-dimensional Measure of Intellectual Humility	22 items; 4 subscales: • Open-mindedness • Intellectual modesty • Corrigibility • Engagement	Not specified	Alfano et al.
Specific Intellectual Humility Scale	9 items; One-factor model	Ranging from $\alpha = .88$ to $\alpha = .96$	Hoyle et al.
General Intellectual Humility Scale	6 items; One-factor model	Ranging from $\alpha = .73$ to $\alpha = .87$	Deffler et al.
Comprehensive Intellectual Humility Scale	22 items; Four-factor structure	Intellect and ego: $\alpha = .88$ Openness to revising one's viewpoint: $\alpha = .87$ Respect for others' viewpoints: $\alpha = .92$ Lack of intellectual overconfidence: $\alpha = .82$	Krumrei-Mancuso & Rouse
Intellectual Humility Scale	16 items; Bi-factorial model	Intellectual openness: $\alpha = .94$ Intellectual arrogance: $\alpha = .93$	McElroy et al.
Biola Intellectual Humility Scale	17 items; Three-factor structure: • Perspective-taking • Low concern for intellectual status • Low intellectual defensiveness	Not specified	Hill et al in Haggard et al.
Relational Humility Scale	16 items; Three-factor structure	Overall: $\alpha = .90$ Global humility: $\alpha = .92$ Superiority: $\alpha = .82$ Accurate view of self: $\alpha = .79$	Davis et al.

From a psychometric perspective, CIHS permits the calculation of an overall score and scores for four distinct subdimensions, reflecting social and epistemic dispositions (Hill et al, 2021; Krumrei-Mancuso et al., 2020). The scale comprises 22 items that measure four core facets, namely (1) Independence of Intellect and Ego, (2) Openness to Revising One's Viewpoint, (3) Respect for Others' Viewpoints, and (4) Lack of Intellectual Overconfidence (Krumrei-Mancuso & Rouse, 2016). CIHS is one of the few instruments designed to capture the multifaceted nature of the construct, providing a more comprehensive theoretical understanding of intellectual humility (Ballantyne, 2023; Gudeliauske, 2020; Huynh & Senger, 2021; Porter et al, 2020).

Previous validation attempts of CIHS in Indonesia have depended on Classical Test Theory (CTT) with Confirmatory Factor Analysis (CFA) and have reported suboptimal reliability estimates (Lubis & Sianipar, 2022). Even though CTT is widely used due to the simplicity (Wang et al, 2024), a key limitation is the strong dependence on sample characteristics (Holland & Hoskens, 2003; Zhang et al, 2023). Item Response Theory (IRT) addresses these limitations by providing detailed information on item functioning, threshold ordering, item difficulty, and measurement precision across different levels of the latent trait. Therefore, IRT-based analyses can more effectively inform decisions regarding scale refinement or revision (Dubbelman et al, 2024; Petrillo et al, 2015).

CFA remains the most frequently used method for scale validation since the concept enables experts to assess whether observed indicators meaningfully relate to underlying latent constructs (Howard et al, 2025). After specifying the model, CFA evaluates whether the estimated parameters adequately reproduce the observed covariance. Model fit indices are used to ascertain the possibility of achieving goals (Howard et al, 2025). This shows that the use of CFA and Rasch analysis is complementary. Evidence-based practice requires scales with known properties, and the knowledge is more complete when CFA and IRT are used (Bean & Bowen, 2021).

The application of Rasch analysis (Wright, 1977) for Indonesian version of CIHS-22 uses the advantages of Rasch model, including specific objectivity, additivity, and parameter separation (Rasch, 1966). These fundamental properties absent in approaches such as CTT enable Rasch model to produce parameter estimates that meet rigorous standards in measurement (Mair & Strasser, 2018). Polytomous models have become available, such as Rating Scale Model (Andrich, 1978), and extend the results of previous studies. The integration of CTT and IRT approaches aims to confirm the theoretical factor structure and evaluate whether the scale meets Rasch measurement requirements, a different and complementary form of internal structure validity evidence. This is achieved by identifying potential item misfits, disordered thresholds, or differential functioning (Cai et

al., 2016; Guo et al., 2024). Therefore, scale functions as a proper measurement instrument, with interpretable factors.

Method

CIHS (Krumrei-Mancuso & Rouse, 2016) comprises 22 items measuring four dimensions, namely Independence of Intellect and Ego (e.g., “I feel small when others disagree with me on topics that are close to my heart”; reverse-coded), Openness to Revising One’s Viewpoint (e.g., “I am willing to change my position on an important issue in the face of good reasons”), Respect for Others’ Viewpoints (e.g., “I can respect others even when I disagree with them on topics that are close to my heart”), and Lack of Intellectual Overconfidence (e.g., “I feel comfortable admitting my intellectual limitations”). All items are rated on a 5-point Likert scale (1 = *strongly disagree* to 5 = *strongly agree*).

This study aims to examine the psychometric properties of the scale using a larger sample and a combined CTT and IRT approach, specifically CFA and Rasch Analysis. Reliability was evaluated through internal consistency (Cronbach’s Alpha (α) & McDonald’s Omega (ω)). Construct validity was assessed using multi-factor CFA by specifying the measurement model *a priori*. In this context, each item was assigned to load on a single hypothesised factor, while correlations among the latent factors were freely estimated. Rasch analysis requires confirmation of unidimensionality of the construct. Despite the

multiple facets, items across subscales measure positions along a single underlying continuum of intellectual humility evaluated through principal component analysis of residuals (PCAR). Reliability was assessed using person and item separation indices to determine the scale discriminative capacity across dimensions, alongside person and item reliability metrics. Furthermore, item fit was examined using the information-weighted fit statistics mean square (infit MNSQ), outlier-sensitive fit statistic mean square (outfit MNSQ), and point-measure correlation (PT-Measure Corr). The residual patterns and item-person fit were analyzed using PCAR and Wright maps, respectively.

All analyses were performed using JASP 0.19.2.0 and Winsteps 5.1.4.0. The data were cleaned first to address invalid and missing values. A total of 504 participants were recruited through Google Forms, and 86 (17.1%) did not complete all measures in the survey. These incomplete responses were excluded from analyses, resulting in a final sample of 418 participants. The sample consisted predominantly of females (305 participants, 72.9%) and undergraduate students (285 participants, 68.1%). The mean age is 25, with participants mainly residing in western Indonesia (357 participants, 85.4%) and having Islam as religion (372 participants, 88.9%). Table 2 provides detailed demographic information, and the data is available online at <https://doi.org/10.6084/m9.figshare.29203418.v1>.

Table 2*Distribution of participant and demographic data*

Sex	<i>n</i>	%
Male	113	27.033
Female	305	72.967
Educational Level	<i>n</i>	%
Undisclosed	2	.478
College/Undergraduate Students	293	70.096
Did not attend college	3	.718
Finished College/Undergraduate	120	28.708
Residency	<i>n</i>	%
West Indonesia	357	85.407
Central Indonesia	35	8.373
East Indonesia	26	6.220

Results

Descriptive analyses were conducted to assess each item response distribution, including mean, standard deviation, skewness, and kurtosis. Likert scale data can be treated as continuous variables when this assumption is satisfied. Skewness values for all items ranged between -1 and 1, indicating that the response distributions were approximately normal (Muthén & Kaplan, 1985).

Most items, specifically C6 to C16, have a high average, indicating that most respondents strongly agree with the statements. Negative skewness and positive kurtosis values show a response distribution that tends to narrow towards agreement. Items such as C7, C8, C11, and C14 show high consistency from respondents, with more than 85% of responses in the two highest categories. In this context, most participants responded to the items homogeneously and positively. C1 to C5 and C17

to C22 show a lower average, ranging from 2.3 to 3.4, with a more spread-out response distribution in the middle and low categories. Skewness of some items is close to zero (C19 and C20), reporting that the respondents do not strongly agree with the statements. Specifically, C20 has the lowest mean (2.32) with a dominant distribution in categories 1 and 2.

Confirmatory Factor Analysis (CFA) Results

The prerequisite for conducting factor analysis was calculating KMO value and performing Bartlett's test of sphericity. The results showed that KMO value for CIHS was .881, which was greater than .7 (Weiner et al, 2017). Bartlett's test of sphericity obtained a chi-square value of $\chi^2 = 3504.762$ with $df = 231$, indicating a statistically significant difference ($p < .001$) and confirming the suitability of the data for factor analysis. CFA was then performed using a 4-factor model in line with a structural model built by Krumrei-Mancuso &

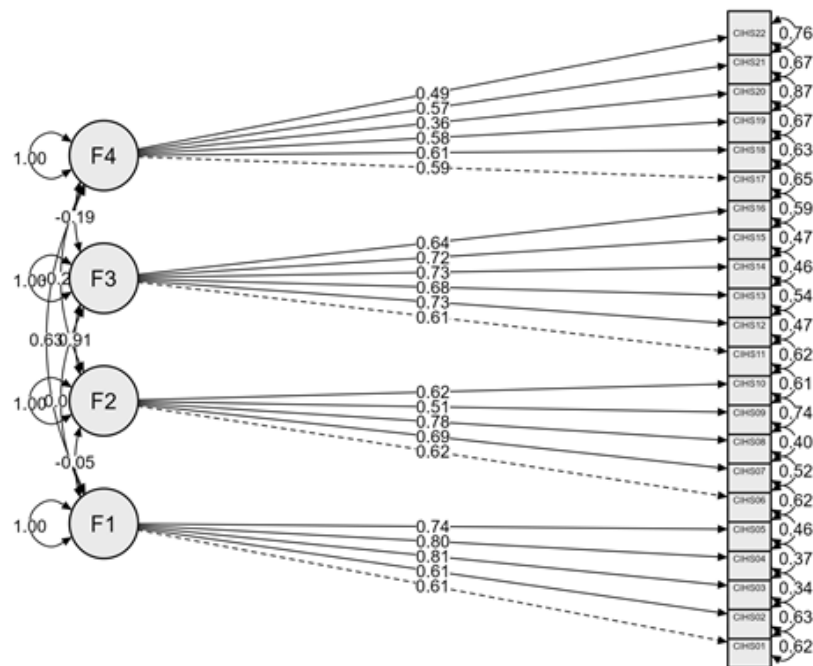
Rouse (2016). Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator was used with robust SE (Beauducel & Herzberg, 2006), listwise deletion to handle missing data, and all scores were standardized. WLSMV estimator was used since the variable was specifically designed for ordinal categorical data (Li, 2016) and recommended for Likert-type scales regardless of the number of response categories (Finney & DiStefano, 2006).

After deriving the estimates (Figure 1), the chi-square for the model is 431.84 ($p=.00$) with Comparative Fit Index (CFI) = .94, Tucker-Lewis Index (TLI) = .93, RMSEA = .05 (CI 90% = .04-.05), Goodness-of-Fit Index (GFI) = .95. A good fit is considered when p in chi-square ($\div 2$) does not reach significance ($p > .05$), CFI is greater than .90, and the root mean square error of approximation (RMSEA) is less than .08 (Brown,

2012; Byrne, 2008). Reliability index for each factor satisfies the acceptable rate (Kumar et al, 2021), Independence of Intellect and Ego ($\hat{\alpha} = .84$, $\hat{\alpha} = .83$, AVE = .51), openness to revise viewpoint ($\hat{\alpha} = .77$, $\hat{\alpha} = .77$, AVE = .40), respect to other viewpoints ($\hat{\alpha} = .83$, $\hat{\alpha} = .84$, AVE = .47), and lack of overconfidence ($\hat{\alpha} = .70$, $\hat{\alpha} = .70$, AVE = .29), with the full-scale coefficient ($\hat{\alpha} = .89$, $\hat{\alpha} = .77$). The sample fits the model and average variance extracted (AVE) values ranged from .29 to .51. AVE values for some dimensions, particularly Lack of Overconfidence, fell below the recommended .50 threshold (Fornell & Larcker, 1981), while composite reliability and omega coefficients exceeded .70, indicating adequate internal consistency despite modest convergent validity (Hair et al, 2010). Some dimensions may exhibit heterogeneous item content and the scale reports acceptable reliability.

Figure 1

Four-factor standardized factor loadings plot of CIHS ($n = 418$)



Variance Explained by Measure

Assumption of unidimensionality of the instrument was tested using Principal Component Analysis Residual (Chou & Wang, 2010; Yang et al, 2023). Unidimensionality was proven when the raw variance explained by the measure is > 30% (Linacre, 2022). This assumption was reported for each aspect of the CIHS, showing that estimating parameters separately (separate calibration) and simultaneously (concurrent calibration) had raw variance explained by the size of > 30%. Approximately 30% of the total variance in item responses was accounted for by the model parameters (Deviana, 2020).

As shown in Table 3, the standardized residual value of the first principal component of each dimension is not greater than 3 (<3.0) (Linacre, 2022). Therefore, the items only measure each aspect intended to represent a large construct, namely CIHS. The results of concurrent calibration showed a score of 6.43%, where the scale was represented by the aspects or multi-dimensional. Results of separate calibration showed that each dimension's unexplained variance was < 3 since the result supported the assumption of unidimensionality. The aspects are also contributing to measuring the same parent construct.

Table 3

Variance of standardized residuals of each dimension

Dimensions	Eigenvalues	Observed (%)
Overall score		
Raw variance explained by the measure	13.81	38.6%
Unexplained variance in the 1st contrast	6.43	18%
Independent of intellect and ego		
Raw variance explained by the measure	6.16	55.2%
Unexplained variance in the 1st contrast	1.54	13.8%
Openness to Revising One's Viewpoint		
Raw variance explained by the measure	4.42	47%
Unexplained variance in the 1st contrast	1.57	16.8%
Respect for Others' Viewpoints		
Raw variance explained by the measure	6.10	50.4%
Unexplained variance in the 1st contrast	1.66	13.7%
Lack of Intellectual Overconfidence		
Raw variance explained by the measure	4.66	43.7%
Unexplained variance in the 1st contrast	1.65	15.5%

Reliability and Separation Indices

In Rasch analysis, reliability is estimated for people and items, presented as coefficients (person separation reliability) and index forms (person and item separation indices). The reliability criteria for the coefficient are $> .70$, indicating that the instrument has good internal consistency. The separation index criteria for people and items used are > 2.5 , which is considered sufficient to conduct a comparative analysis at the group level (Deviana, 2020; Tennant & Conaghan, 2007).

The person separation coefficient is not fulfilled for all dimensions (1.51-2.02). The most challenging item is C20, "In important things, I will not be influenced by other people's points of view," which translates into "Dalam hal-hal penting, saya tidak akan terpengaruh

oleh sudut pandang orang lain". The easiest item to endorse is C11, "I respect others even when I disagree with them on important issues," which translates into "Saya tetap menghormati orang lain bahkan ketika saya tidak sependapat dengan mereka tentang isu-isu penting."

Table 4 shows that the person reliability estimate on the instrument is approximately .70 (range .71 - .80). Therefore, each dimension has good internal consistency and can differentiate individuals based on the level of attributes measured. The instrument can reliably differentiate individuals with different levels of intellectual humility. This result is followed by the item reliability from the four dimensions (range = .96 - .99). Lack of Intellectual Overconfidence obtained the highest item reliability of 0.99.

Table 4

Person & item reliability indices

Dimensions	Person Reliability	Person Separation	Item Reliability	Item Separation
Independent of intellect and ego	.80	2.02	.97	5.89
Openness to Revising One's Viewpoint	.71	1.57	.98	7.49
Respect for Others' Viewpoints	.76	1.80	.96	4.85
Lack of Intellectual Overconfidence	.70	1.51	.99	9.05

Fit of the Items

The ideal criteria for the fit of item analysis are infit MNSQ and outfit MNSQ values between .6 and 1.4 (Rakami et al., 2020), suggesting a good fit to the model. The analysis showed the infit and outfit MNSQ values for

each item of the scale ranged from .74 to 1.30 and .69 to 1.37, respectively. The correlation coefficient measures how closely the items are consistent with the measurement target, at a minimum acceptable value of .5 (Dong et al., 2020). In this study, PT-Measure Corr values

ranged from .5 to .81, exceeding the minimum reference value. Therefore, the scale items are

consistent with the measurement target as reported in Table 5.

Table 5

Analysis of CIHS items

Items	Item measure	INFIT MNSQ	OUTFIT MNSQ	PT-Measure Corr.
C1	.55	1.13	1.10	.72
C2	.45	1.27	1.31	.70
C3	-.48	.93	.90	.78
C4	-.45	.78	.80	.80
C5	-.07	.85	.82	.81
C6	-.27	1.03	1.06	.66
C7	-.85	.84	.83	.71
C8	-.54	.74	.69	.78
C9	1.05	1.30	1.37	.69
C10	.61	1.02	1.11	.70
C11	-.43	1.13	1.10	.70
C12	-.25	.86	.84	.78
C13	1.04	1.11	1.15	.70
C14	-.35	.86	.79	.77
C15	-.21	.98	.98	.73
C16	.19	1.06	1.02	.71
C17	-.11	.76	.78	.67
C18	-.45	.88	.88	.65
C19	.44	1.00	1.06	.58
C20	1.03	1.23	1.22	.50
C21	-.78	1.08	1.07	.69
C22	-.13	1.02	1.02	.64

Item-Person Matching

The subsequent analysis is reported through Wright Map, which converts the original scores of individual ability and item difficulties into logit values on the same scale, depicting the suitability of the items for the individual (Li et al, 2025; Pichardo et al, 2018). This map serves as the leading indicator of the overall quality of

the scale and simultaneously shows the ability level of respondents who participated in filling out the instrument and the difficulty level of all items. The results of the analysis of each dimension of intellectual humility on the instrument are shown in Figure 2.

The means on both sides should be close to zero, with a difference of less than 1 logit (Figure 2). A gap that exceeds one logit shows a

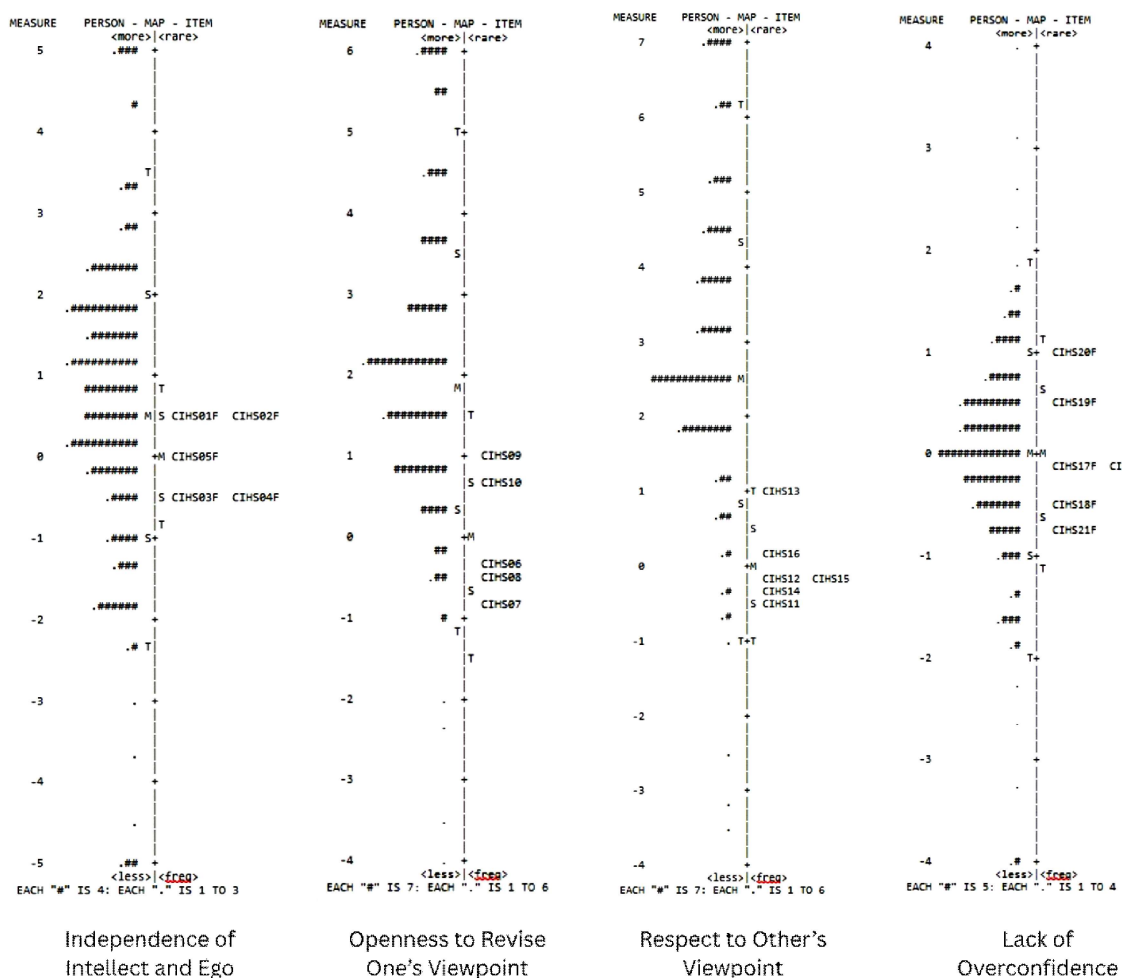
discrepancy between individual ability and item difficulty level. The symbols “S” and “T” represent one and two standard deviations, respectively. On the left side of the map, individual ability is shown, with higher positions indicating higher levels. In contrast, the right side depicts item difficulty, with higher positions indicating more difficult items to endorse (Raykov et al, 2020).

Figure 2 shows a reasonably good balance between participant ability and item difficulty in the four dimensions, with a relative

distribution. For the Openness to Revise One’s Viewpoint and Respect for Others’ Viewpoint dimensions, the distribution of individual ability is above the item. Therefore, the items are relatively easy for respondents to agree with. The average difference between ability and item difficulty remains below one logit threshold. The distribution of participant ability shows regular pattern, with most individuals in the middle ability range. Most items in the scale are also in the same difficulty range since the scale is representative of the samples.

Figure 2

Wright’s person-item map for the CIHS-22 items



Response Category Fit

CIHS includes five response categories, from “Strongly Does Not Apply to Me” to “Strongly Applies to Me”. Table 6 reports the percentage of responses in each category, the estimated difficulty level, and the mean squared value (MNSQ) to evaluate the appropriateness and reliability. The table shows that the difficulty

levels of the response categories are in ascending order as expected, as well as all response categories with MNSQ values in the acceptable range of .5–1.5 (Linacre, 2022; Rakami et al, 2020). Infit and Outfit values below .5 report that the item is overfit (too ideal). Meanwhile, Infit and Outfit values above 1.5 are considered to be underfit or inconsistent (Linacre, 2022).

Table 6

Rating scale diagnostics for each dimension

	Category	Threshold	Observed Count (%)	Observed Average	Infit MNSQ	Outfit MNSQ
Independent of intellect and ego	Strongly Does Not Apply to Me	None	131 (6)	-1.36	2.21	1.90
	Does Not Apply to Me	-3.07	455 (22)	-.92	.88	.90
	Neutral	-.39	561 (27)	.26	.76	.77
	Applies to Me	.73	634 (30)	1.41	.75	.73
	Strongly Applies to Me	2.74	309 (15)	2.27	1.07	1.12
Openness to Revising One's Viewpoint	Strongly Does Not Apply to Me	None	23 (1)	-1.10	1.65	2.21
	Does Not Apply to Me	-2.40	63 (3)	-.11	1.39	1.66
	Neutral	-1.86	414 (20)	.42	.90	.86
	Applies to Me	.31	1082 (52)	2.00	.79	.82
	Strongly Applies to Me	3.95	508 (24)	3.81	1.00	.95
Respect for Others' Viewpoints	Strongly Does Not Apply to Me	None	17 (1)	-2.05	1.57	2.03
	Does Not Apply to Me	-2.71	44 (2)	-.49	1.40	1.56
	Neutral	-2.22	389 (16)	.76	1.01	1.00
	Applies to Me	.46	1351 (54)	2.53	.81	.81
	Strongly Applies to Me	4.47	707 (28)	4.50	.99	.92
Lack of Intellectual Overconfidence	Strongly Does Not Apply to Me	None	249 (10)	-1.40	1.39	1.36
	Does Not Apply to Me	-2.14	558 (22)	-.89	.73	.75
	Neutral	-.87	979 (39)	.07	.83	.84
	Applies to Me	.96	533 (21)	.75	.85	.85
	Strongly Applies to Me	2.04	189 (8)	1.21	1.20	1.26

Note. Threshold: Rasch-Andrich thresholds: The relationship between adjacent categories corresponds to the point where the probability curves of adjacent categories intersect; Observed count (%): the number of occurrences of a category; Observed average: the average ability of a person to answer each category; Infit & Outfit: a mean square statistic. The expected value is 1. A value of .5-1.5 is a good measure.

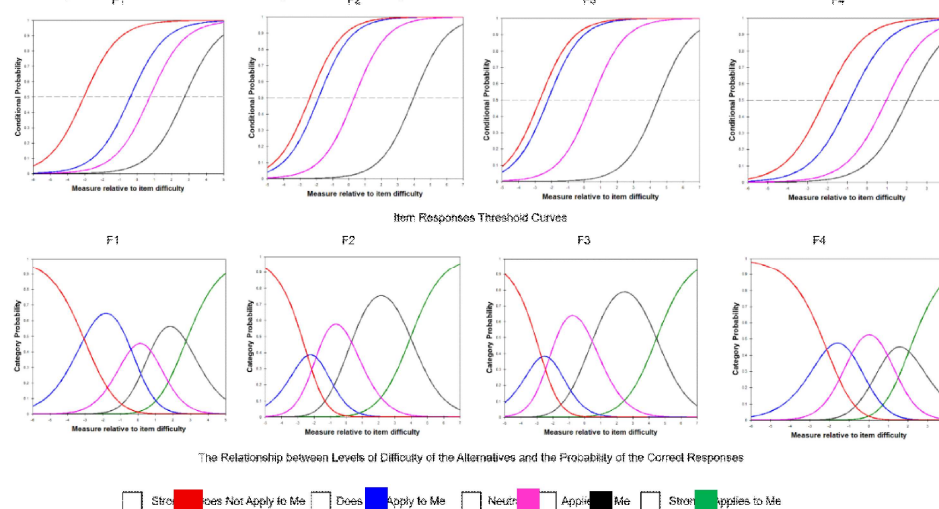
Independence of Intellect and Ego, Openness to Revise One's Opinion, and Respect for Others' Viewpoint had elevated Infit and Outfit MNSQ values in "Strongly Does Not Apply to Me" response category. This reflects sparse observations and instability at the extreme ends of the latent trait rather than substantive category misfit. Extreme categories are known to exhibit inflated fit indices and should be interpreted cautiously (Linacre, 2022). Meanwhile, Lack of Overconfidence has Infit and Outfit MNSQ values in reasonable limits in all categories, including "Strongly Does Not Apply to Me". The responses in this category tend to be appropriate and consistent. In Openness to Revise One's Opinion, the Outfit MNSQ value for the "Does Not Apply to Me" category was also identified as exceeding the upper limit, indicating an aberrant response. Possible causes include participants with high ability levels giving inconsistent responses to

easy items. This inconsistency can report misconceptions, confusing item interpretations, or unusual response behavior such as guessing (Liu et al., 2019).

The increasing difficulty of the response options can be seen in Figure 3, where the scale effectively reflected the varying ability levels of the respondents. As the ability level increased, the probability of selecting more difficult options increased. In this context, the most difficult option ("Strongly Applies to Me") becomes the correct answer since the scale can effectively differentiate between different trait levels. These results show that CIHS effectively distinguishes between varying levels of ability and item difficulty. The peak values of the four-dimensional item characteristic curves were all different from each other and in the same order. Therefore, there is a reasonable degree of differentiation among the dimensional items.

Figure 3

The item characteristic curve of the four dimensions of CIHS, F1 = Independent of Intellect and Ego, F2 = Openness to Revise One's Opinion, F3 = Respect for Others' Viewpoint, F4 = Lack of Overconfidence



Discussion

The main objective was to explore the psychometric properties of CIHS used in Indonesian population. CTT approach was used with CFA, and IRT approach with Rasch Rating Scale Model. The results showed that the structural model of CIHS was adequate and reliable with Indonesian sample. The good fit indices of CFA showed the psychometric solidity of the 22-item structure of Indonesian version of CIHS. The reliability coefficient supported the internal consistency of the measure. The outcomes suggested that Indonesian version of the scale was a reliable tool and manifested roughly similar psychometric properties to the original. Therefore, CIHS, in the four aspects of Independence of Intellect and Ego, Openness to Revising One's Viewpoint, Respect for Others' Viewpoints, and Lack of Intellectual Overconfidence, assessed intellectual humility.

The relatively low AVE score for the Lack of Overconfidence (.29) compared to the other dimensions suggested that the scale showed acceptable reliability, even though the dimensions reported heterogeneous item content as indicated by the sufficient value of alpha (α) and omega (ω) coefficients. Previous studies also experienced similar results where the values of lack of overconfidence were lower than the other scales (Kroplewski et al, 2022; Stankov & Lee, 2014; Zhdanova & Shchebetenko, 2024). According to Kroplewski et al (2022), this result may be attributable to

differences in individuals' perceptions of overconfidence. In different regions, the term has been analyzed in distinct ways, often leading to inconsistent results (Baek & Cho, 2022; Moore & Healy, 2008).

The first limitation of this study is the distribution of respondents that mostly comprised of muslim and resided in the western part of Indonesia. The inequality of gender, occupation, and educational level can also highly impact the results (Fisher, 1991). The second limitation is the use of real-world samples to provide psychometric properties of scales. The use of a simulation study (Monte-Carlo Simulation; see Nuel et al., 2024) can present more accurate dynamics of the scale, particularly the psychological scale (Gronier, 2022; Nuel et al, 2024). Moreover, the use of self-reporting to measure intellectual humility may be questioned. Other measures, such as the informant-reported scale (Meagher, 2022), are used to holistically assess the construct. Differences in the person separation coefficient, which influenced infit and MNSQ statistics (Mousavi & Cui, 2020), may be reduced by rewording the items to minimize normative wording and susceptibility to social desirability bias.

Conclusions

In conclusion, the scale developed in this study adds to the existing body of intellectual humility measures, expanding the range of available instruments for assessing the

construct CIHS was investigated using CFA and Rating Scale Model (IRT). CTT analysis reported acceptable reliability and good model fit, while Rating Scale Model confirmed the good fit and measurement precision by further expansion. The combination of the approaches ensured the scale robustness in assessing intellectual humility, particularly in Indonesian population. Future studies should test the scale in a diverse population to establish its applicability and generalizability.

Suggestion

Future studies could expand the samples to larger regions and to various roles in society. The construct of lack of overconfidence should be empirically examined. In addition, further scale refinement is recommended through the development of items representing more extreme levels of difficulty at both the higher and lower ends of the continuum, as well as by achieving a more balanced distribution of item difficulty. These refinements would enable the scale to more comprehensively capture the full range of participant abilities. Other IRT models to be considered include Samejima's Graded Response Model (GRM) and Multidimensional Item Response Theory (MIRT).

References

- Alfano, M., Iurino, K., Stey, P., Robinson, B., Christen, M., Yu, F., & Lapsley, D. (2017). Development and validation of a multi-dimensional measure of intellectual humility. *PloS One*, 12(8), e0182950. <https://doi.org/10.1371/journal.pone.0182950>
- Andrich, David. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2(4), 581–594. <https://doi.org/10.1177/014662167800200413>
- Baek, H. Y., & Cho, D. D. (2022). Overconfidence and risky investment choices Hyungkee. *Economics Bulletin, AccessEcon*, 42(4), 2267–2278. <https://ideas.repec.org/a/ebl/ecbull/eb-22-00264.html>
- Ballantyne, N. (2023). Recent work on intellectual humility: A philosopher's perspective. *The Journal of Positive Psychology*, 18(2), 200–220. <https://doi.org/10.1080/17439760.2021.1940252>
- Barrett, J. L. (2017). Intellectual humility. *The Journal of Positive Psychology*, 12(1), 1–2. <https://doi.org/10.1080/17439760.2016.1167945>
- Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in cfa. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2
- Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item response theory. In *annual review of statistics and its application* (Vol. 3, pp. 297–321). Annual Reviews Inc. <https://doi.org/10.1146/annurev-statistics-041715-033702>
- Chancellor, J., & Lyubomirsky, S. (2013). Humble beginnings: Current trends, state perspectives, and hallmarks of humility. *Social and Personality Psychology Compass*, 7(11), 819–833. <https://doi.org/https://doi.org/10.1111/spc3.12069>
- Chou, Y.-T., & Wang, W.-C. (2010). Checking dimensionality in item response models with principal component analysis on standardized residuals. *Educational and*

- Psychological Measurement*, 70(5), 717–731. <https://doi.org/10.1177/0013164410379322>
- Davis, D. E., Rice, K., McElroy, S., DeBlaere, C., Choe, E., Van Tongeren, D. R., & Hook, J. N. (2016). Distinguishing intellectual humility and general humility. *The Journal of Positive Psychology*, 11(3), 215–224. <https://doi.org/10.1080/17439760.2015.1048818>
- Deviana, T. (2020). *Penerapan rasch model dan response times modeling untuk mendeteksi aberrant response behavior*. <https://repository.uinjkt.ac.id/dspace/bitstream/123456789/79277/1/TINA-DEVIANA-FPSI.pdf>
- Dong, D., Jin, J., Oerlemans, S., Yu, S., Yang, S., Zhu, J., & Xu, R. H. (2020). Validation of the Chinese EORTC chronic lymphocytic leukaemia module - application of classical test theory and item response theory. *Health and Quality of Life Outcomes*, 18(1), 96. <https://doi.org/10.1186/s12955-020-01341-z>
- Dubbelman, M. A., Postema, M. C., Jutten, R. J., Harrison, J. E., Ritchie, C. W., Aleman, A., de Jong, F. J., Schalet, B. D., Terwee, C. B., van der Flier, W. M., Scheltens, P., & Sikkes, S. A. M. (2024). What's in a score: A longitudinal investigation of scores based on item response theory and classical test theory for the Amsterdam Instrumental Activities of Daily Living Questionnaire in cognitively normal and impaired older adults. *Neuropsychology*, 38(1), 96–105. <https://doi.org/10.1037/neu0000914>
- Finney, S., & DiStefano, C. (2006). *Nonnormal and categorical data in structural equation models* (pp. 269–314). <https://doi.org/10.1108/978-1-62396-246-320251015>
- Fisher, W. P. (1991). The rasch debate: validity and revolution in educational measurement. <https://eric.ed.gov/?id=ED331857>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.2307/3151312>
- Gronier, G. (2022). Psychometric analyses in the transcultural adaptation of psychological scales. In S. Misciagna (Ed.), *S. Misciagna (Ed.) Psychometrics - new insights* (pp. 1–13). IntechOpen. <https://doi.org/10.5772/intechopen.105841>
- Gudeliauske, G. (2020). Intellectual humility as strenght in the processes of the formation and development of knowledge. *SOCIETY. INTEGRATION. EDUCATION. Proceedings of the International Scientific Conference*, 3, 187. <https://doi.org/10.17770/sie2020vol3.4928>
- Guo, Zhichen, Wang, Daxun, Cai, Yan, & Tu, Dongbo. (2024). An item response theory model for incorporating response times in forced-choice measures. *Educational and Psychological Measurement*, 84(3), 450–480. <https://doi.org/10.1177/00131644231171193>
- Haggard, M., Rowatt, W. C., Leman, J. C., Meagher, B., Moore, C., Fergus, T., Whitcomb, D., Battaly, H., Baehr, J., & Howard-Snyder, D. (2018). Finding middle ground between intellectual arrogance and intellectual servility: Development and assessment of the limitations-owning intellectual humility scale. *Personality and Individual Differences*, 124, 184–193. <https://doi.org/https://doi.org/10.1016/j.paid.2017.12.014>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Hill, P. C., Lewis Hall, M. E., Wang, D., & Decker, L. A. (2021). Theistic intellectual humility and well-being: does ideological context matter? *The Journal of Positive Psychology*, 16(2), 155–167. <https://doi.org/10.1080/17439760.2021.1911111>

- doi.org/10.1080/17439760.2019.1689424
- Holland, P., & Hoskens, M. (2003). Classical test theory as a first-order item response theory: Application to true-score prediction from a possibly nonparallel test. *Psychometrika*, 68, 123–149. <https://doi.org/10.1007/BF02296657>
- Hook, J. N., Farrell, J. E., Johnson, K. A., Van Tongeren, D. R., Davis, D. E., & Aten, J. D. (2017). Intellectual humility and religious tolerance. *The Journal of Positive Psychology*, 12(1), 29–35. <https://doi.org/10.1080/17439760.2016.1167937>
- Hoyle, R. H., Davisson, E. K., Diebels, K. J., & Leary, M. R. (2016). Holding specific views with humility: Conceptualization and measurement of specific intellectual humility. *Personality and Individual Differences*, 97, 165–172. <https://doi.org/https://doi.org/10.1016/j.paid.2016.03.043>
- Howard, M. C., Boudreaux, M., Cogswell, J., Manix, K. G., & Oglesby, M. T. (2025). A literature review of model fit and model comparisons with confirmatory factor analysis: Formalizing the informal in organizational science. *Applied Psychology*, 74(1), e12592. <https://doi.org/10.1111/apps.12592>
- Huynh, H. P., & Senger, A. R. (2021). A little shot of humility: Intellectual humility predicts vaccination attitudes and intention to vaccinate against COVID-19. *Journal of Applied Social Psychology*, 51(4), 449–460. <https://doi.org/https://doi.org/10.1111/jasp.12747>
- Kroplewski, Z., Krumrei-Mancuso, E. J., Bielecka, G., & Szczeciński, M. (2022). A preliminary validation of the polish version of the comprehensive intellectual humility scale (cihs). *Psychology Research and Behavior Management*, 15(null), 3627–3638. <https://doi.org/10.2147/PRBM.S380470>
- Krumrei-Mancuso, E. J., Haggard, M. C., LaBouff, J. P., & Rowatt, W. C. (2020). Links between intellectual humility and acquiring knowledge. *The Journal of Positive Psychology*, 15(2), 155–170. <https://doi.org/10.1080/17439760.2019.1579359>
- Krumrei-Mancuso, E. J., & Rouse, S. V. (2016). The development and validation of the comprehensive intellectual humility scale. *Journal of Personality Assessment*, 98(2), 209–221. <https://doi.org/10.1080/00223891.2015.1068174>
- Krumrei Mancuso, E. (2017). Intellectual humility and prosocial values: Direct and mediated effects. *The Journal of Positive Psychology*, 12(1), 1–16. <https://doi.org/10.1080/17439760.2016.1167938>
- Kumar, A. P., Omprakash, A., Mani, P. K. C., Swaminathan, N., Maheshkumar, K., Maruthy, K. N., Sathiyasekaran, B. W. C., Vijayaraghavan, P. V., & Padmavathi, R. (2021). Validation of internal structure of self-directed learning readiness scale among indian medical students using factor analysis and the structural equation modelling approach. *BMC Medical Education*, 21(1), 614. <https://doi.org/10.1186/s12909-021-03035-6>
- Leary, Mark R, Diebels, Kate J, Davisson, Erin K, Jongman-Sereno, Katrina P, Isherwood, Jennifer C, Raimi, Kaitlin T, Deffler, Samantha A, & Hoyle, Rick H. (2017). Cognitive and interpersonal features of intellectual humility. *Personality and Social Psychology Bulletin*, 43(6), 793–813. <https://doi.org/10.1177/0146167217697695>
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Li, C., Lin, Y., Tosun, B., Wang, P., Guo, H. Y., Ling, C. R., Qi, R., Luo, Q. Y., Wang, Y., Huang, F.,

- Wang, J., Ma, S. H., Xu, D. F., Wu, S. Z., & Zhang, L. (2025). Psychometric evaluation of the Chinese version of the BENEFITS-CCCSAT based on CTT and IRT: a cross-sectional design translation and validation study. *Frontiers in Public Health*, Volume 13-2025. <https://doi.org/10.3389/fpubh.2025.1532709>
- Linacre, J. M. (2022). R Statistics: survey and review of packages for the estimation of Rasch models. In *International journal of medical education* (Vol. 13, pp. 171–175). <https://doi.org/10.5116/ijme.629dd88f>
- Liu, T., Sun, Y., Li, Z., & Xin, T. (2019). The impact of aberrant response on reliability and validity. *Measurement: Interdisciplinary Research and Perspectives*, 17(3), 133–142. <https://doi.org/10.1080/15366367.2019.1584848>
- Lubis, S. I., & Sianipar, A. (2022). How religious tolerance can emerge among religious people: An investigation on the roles of intellectual humility, cognitive flexibility, and trait aggressiveness. *Asian Journal of Social Psychology*, 25(2), 276–287. <https://doi.org/https://doi.org/10.1111/ajsp.12493>
- Mair, P., & Strasser, H. (2018). Large-scale estimation in Rasch models: asymptotic results, approximations of the variance-covariance matrix of item parameters, and divide-and-conquer estimation. *Behavior metrika*, 45(1), 189–209. <https://doi.org/10.1007/s41237-018-0046-z>
- McElroy, Stacey E, Rice, Kenneth G, Davis, Don E, Hook, Joshua N, Hill, Peter C, Worthington Jr., Everett L, & Van Tongeren, Daryl R. (2014). Intellectual humility: Scale Development and theoretical elaborations in the context of religious leadership. *Journal of Psychology and Theology*, 42(1), 19–30. <https://doi.org/10.1177/009164711404200103>
- Meagher, B. R. (2022). An assessment of self and informant data for measuring intellectual humility. *Personality and Individual Differences*, 184, 111218. <https://doi.org/https://doi.org/10.1016/j.paid.2021.111218>
- Moore, D., & Healy, P. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502–517. <https://doi.org/10.1037/0033-295X.115.2.502>
- Mousavi, A., & Cui, Y. (2020). The effect of person misfit on item parameter estimation and classification accuracy: a simulation study. In *Education Sciences*, 10 (11), 324 <https://doi.org/10.3390/educsci10110324>
- Muthén, B., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38(2), 171–189. <https://doi.org/https://doi.org/10.1111/j.2044-8317.1985.tb00832.x>
- Nuel, I., Nurra, C., Dujols, O., Tagand, M., Webster, G., Gorgo, G., & Muller, D. (2024). “I know that i don’t know everything”: Validation of the general intellectual humility scale in french. *Collabra: Psychology*, 10(1), 125126. <https://doi.org/10.1525/collabra.125126>
- Petrillo, J., Cano, S. J., McLeod, L. D., & Coon, C. D. (2015). Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value in Health/ : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, 18(1), 25–34. <https://doi.org/10.1016/j.jval.2014.10.005>
- Pichardo, M. C., Cano, F., Garzón-Umerenkova, A., de la Fuente, J., Peralta-Sánchez, F. J., & Amate-Romera, J. (2018). Self-regulation questionnaire (srq) in spanish adolescents: Factor structure and rasch analysis. *Frontiers in Psychology*, Volume 9-2018. <https://doi.org/10.3389/fpsyg.2018.01370>
- Porter, T., Baldwin, C. R., Warren, M. T., Murray, E. D., Cotton Bronk, K., Forgeard, M. J. C., Snow, N. E., & Jayawickreme, E. (2022). Clarifying the content of intellectual humility: a systematic review and

- integrative framework. *Journal of Personality Assessment*, 104(5), 573–585. <https://doi.org/10.1080/00223891.2021.1975725>
- Porter, T., Elnakouri, A., Meyers, E. A., Shibayama, T., Jayawickreme, E., & Grossmann, I. (2022). Predictors and consequences of intellectual humility. *Nature Reviews Psychology*, 1(9), 524–536. <https://doi.org/10.1038/s44159-022-00081-9>
- Porter, T., & Schumann, K. (2018). Intellectual humility and openness to the opposing view. *Self and Identity*, 17(2), 139–162. <https://doi.org/10.1080/15298868.2017.1361861>
- Porter, T., Schumann, K., Selmezy, D., & Trzesniewski, K. (2020). Intellectual humility predicts mastery behaviors when learning. *Learning and Individual Differences*, 80, 101888. <https://doi.org/https://doi.org/10.1016/j.lindif.2020.101888>
- Rakami, N. M. H. N., Ismail, N. A. H., Abu Kassim, N. L., & Idrus, F. (2020). Validation of egalitarian education questionnaire using rasch measurement Model. *Journal of Applied Measurement*, 21(1), 91–100.
- Rasch, G. (1966). An individualistic approach to item analysis. In *Readings in mathematical social scienc* (pp. 89–107). <https://rasch.org/memo19662.pdf>
- Raykov, T., Marcoulides, G. A., & Huber, C. (2020). Inaccurate individual ability estimates with three-parameter item response models in mixture settings. *Measurement: Interdisciplinary Research and Perspectives*, 18(1), 15–22. <https://doi.org/10.1080/15366367.2019.1641876>
- Stankov, Lazar, & Lee, Jihyun. (2014). Overconfidence across world regions. *Journal of Cross-Cultural Psychology*, 45(5), 821–837. <https://doi.org/10.1177/0022022114527345>
- Tangney, J. P. (2000). Humility: Theoretical perspectives, empirical findings and directions for future research. *Journal of Social and Clinical Psychology*, 19(1), 70–82. <https://doi.org/10.1521/jscp.2000.19.1.70>
- Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis and Rheumatism*, 57(8), 1358–1362. <https://doi.org/10.1002/art.23108>
- Van Tongeren, D. R., Hook, J. N., Ramos, M. J., Edwards, M., Worthington Jr, E. L., Davis, D. E., Ruiz, J. M., Reid, C. A., Garthe, R. C., Nonterah, C. W., Cowden, R. G., Opare-Henaku, A., Connelly, R., Omoruyi, O., Nkomo, T. S., & Osa-Larbi, J. A. (2019). The complementarity of humility hypothesis: Individual, relational, and physiological effects of mutually humble partners. *The Journal of Positive Psychology*, 14(2), 178–187. <https://doi.org/10.1080/17439760.2017.1388433>
- Wang, Y.-Y., Tan, X.-W., Hu, Z.-N., & Zhong, L.-L. (2024). Development and validation of a multidimensional perceived media convergence scale: A classical test theory and item response theory approach. *IEEE Access*, PP, 1. <https://doi.org/10.1109/ACCESS.2024.3383141>
- Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science*, 12(1), 108. <https://doi.org/10.1186/s13012-017-0635-3>
- Wright, B. D. (1977). Solving measurement problems with the rasch model. *Journal of Educational Measurement*, 14(2), 97–116. <https://doi.org/https://doi.org/10.1111/j.1745-3984.1977.tb00031.x>

- Yang, Y., Peng, Y., Li, W., Lu, S., Wang, C., Chen, S., & Zhong, J. (2023). Psychometric evaluation of the academic involution scale for college students in China: An application of Rasch analysis. *Frontiers in Psychology, Volume 14-2023*. <https://doi.org/10.3389/fpsyg.2023.1135658>
- Zachry, C. E., Phan, L. V., Blackie, L. E. R., & Jayawickreme, E. (2018). Situation-based contingencies underlying wisdom-content manifestations: Examining intellectual humility in daily life. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 73(8), 1404–1415. <https://doi.org/10.1093/geronb/gby016>
- Zhang, D., Wang, C., Yuan, T., Li, X., Yang, L., Huang, A., Li, J., Liu, M., Lei, Y., Sun, L., Zhang, J., & Zhang, L. (2023). Psychometric properties of the coronavirus anxiety scale based on classical test theory (ctt) and item response theory (irt) models among chinese front-line healthcare workers. *BMC Psychology*, 11(1), 224. <https://doi.org/10.1186/s40359-023-01251-x>
- Zhdanova, Y. A., & Shchebetenko, S. A. (2024). Psychometric properties of a russian version of the comprehensive scale of intellectual humility. *National Psychological Journal*, 19(2), 131–142. <https://doi.org/https://doi.org/10.11621/npj.2024.0211>
- Zmigrod, L., Zmigrod, S., Rentfrow, P. J., & Robbins, T. W. (2019). The psychological roots of intellectual humility: The role of intelligence and cognitive flexibility. *Personality and Individual Differences*, 141, 200–208. <https://doi.org/https://doi.org/10.1016/j.paid.2019.01.016>
- ✱
- Received 5 November 2025
Revised 16 January 2026
Accepted 30 January 2026