



Contents lists available at [Journal ELORA](#)

JRTI (Jurnal Riset Tindakan Indonesia)

ISSN: 2502-079X (Print) ISSN: 2503-1619 (Electronic)

Journal homepage: <https://jrti.eloracenter.org/jrti>



Psychometric evaluation of MIND-RISK for assessing mental health risk among university students

Parid Rilo Pambudi^{*1}, Agustin Mega Purnamawati², Novia Damayanti³, Hana Athia Akhzalini⁴, Sri Retno Wulandari⁵, Revi Aprilia Fitri⁶, Nining Maizura⁷, Arifah Wulandari⁸

^{1, 2, 3, 4, 5, 6} Universitas Islam Balitar, Indonesia

⁷ Universitas Negeri Malang, Indonesia

⁸ Universitas Sebelas Maret, Indonesia

Article Info

Article history:

Received Jul 13th, 2026

Revised Aug 30th, 2026

Accepted Sept 28th, 2026

Keyword:

MIND-RISK,
Mental health risk,
University students,
Psychometric evaluation,
Rasch model,
Psychological assessment

ABSTRACT

This study evaluated the psychometric characteristics of the Mental Health INdex for Academic Distress and Psychological RISK (MIND-RISK), a multidomain instrument designed to assess both psychological risk and protective factors among university students. Data from 100 students were analyzed using the Rasch Partial Credit Model across six constructs: Academic Stress Load, Intolerance of Uncertainty, Academic Adaptability, Emotion Regulation, Psychological Resilience, and Psychological Flourishing. The results showed acceptable item fit for all 48 items, reliable measurement for both the Risk Factor and Protective Factor Indices, and properly functioning five-point response categories. Although lower reliability was found for Intolerance of Uncertainty and some residual dependence was identified, the full 48-item structure was retained. Overall, MIND-RISK demonstrates promising psychometric properties and may be useful for preliminary psychological assessment and profile mapping in university guidance and counseling services. Further validation with larger and more diverse samples and additional psychological measures is recommended before wider standardized use.



© 2026 The Authors.

This is an open access article under the CC BY-NC-SA license

(<https://creativecommons.org/licenses/by-nc-sa/4.0>)

Corresponding Author:

Parid Rilo Pambudi,

Universitas Islam Balitar

Email: paridrilopambudi@unisbablitar.ac.id

Introduction

Mental health among university students has become an important concern in higher education because psychological difficulties may interfere with students' academic functioning, persistence, and broader personal development. University students encounter multiple and overlapping demands, including academic workload, performance expectations, social adjustment, uncertainty about the future, and changes associated with the transition into adulthood. These challenges may be particularly salient among first-year students who must simultaneously adapt to higher education and changes in their living environment, with academic anxiety emerging as an important psychological concern during this transition (Maizura et al., 2026). Empirical evidence indicates that university students experience substantial levels of anxiety, depression, stress, and other psychological difficulties, with academic stress and intolerance of uncertainty emerging as important factors associated with poorer mental health outcomes (Ahmed et al., 2023; Worsley et al., 2024). Mental health

difficulties may also have consequences for students' educational trajectories, including academic functioning and university continuation (Leow et al., 2025; Ramos-Monsivais et al., 2024). These conditions underscore the importance of assessment approaches that can identify students' mental health risks before psychological difficulties become more severe or interfere substantially with academic functioning.

The Indonesian context further emphasizes the importance of early identification of psychological difficulties in higher education. Mental health screening among Indonesian university students has identified psychological difficulties that may require counseling, stress-management support, and other preventive services (Astuti et al., 2024). These findings provide direct evidence that mental health assessment is relevant within Indonesian higher education settings. At an earlier developmental stage, the Indonesia–National Adolescent Mental Health Survey (I-NAMHS) reported substantial mental health problems among Indonesian adolescents aged 10–17 years (Wahdi et al., 2023). Although the I-NAMHS population differs from university students, these findings provide broader developmental context indicating that psychological vulnerability may already emerge before entry into higher education. Within university populations, academic stress, intolerance of uncertainty, and other psychological factors have also been associated with anxiety, depression, and impaired academic functioning (Ramos-Monsivais et al., 2024; Worsley et al., 2024). Taken together, these findings support the need for assessment approaches that are sensitive not only to psychological symptoms but also to academic vulnerability and psychological resources relevant to university life.

Students' mental health is shaped by the interaction between risk and protective factors rather than by psychological symptoms alone. Academic stress represents an important contextual risk factor because excessive academic demands, workload, and performance expectations may contribute to psychological distress (Quincho et al., 2021). Intolerance of uncertainty may constitute another vulnerability factor, particularly when students face ambiguous academic, social, or future-oriented situations; research among university students has demonstrated associations between intolerance of uncertainty and anxiety, depression, and other indicators of psychological distress (Charbonnier et al., 2023; Huang et al., 2023). Conversely, adaptive psychological resources may protect students from the adverse effects of stress. Emotion regulation and academic adjustment have been identified as important factors associated with students' ability to manage academic demands, while resilience represents an important capacity for coping with adversity and maintaining psychological functioning (Connor & Davidson, 2003). Psychological flourishing provides a complementary perspective by emphasizing positive functioning and well-being rather than the absence of psychological symptoms alone (Keyes, 2002). Meaning-making is also relevant to university students' psychological functioning, particularly when they encounter developmental uncertainty and quarter-life challenges, as the ability to construct personal meaning may become an important psychological resource (Akhzalini et al., 2026). Taken together, these perspectives suggest that a meaningful assessment of student mental health should consider both vulnerability and the psychological resources that may enable students to cope adaptively with academic and life demands.

Based on this perspective, the present study developed the *Mental Health INDEX for Academic Distress and Psychological RISK* (MIND-RISK) as an integrative assessment framework for university students. MIND-RISK consists of two complementary dimensions: the *Risk Factor Index*, comprising *academic stress load* and *intolerance of uncertainty*, and the *Protective Factor Index*, comprising *academic adaptability*, *emotion regulation*, *psychological resilience*, and *psychological flourishing*. This structure is intended to provide a more comprehensive representation of students' mental health profiles by examining psychological vulnerability alongside available protective resources. Rather than conceptualizing mental health solely as the presence or absence of psychopathological symptoms, MIND-RISK adopts a multidimensional perspective in which students may differ in both their exposure to psychological risk and their capacity to cope with academic and psychosocial demands.

In the present study, mental health risk is operationally conceptualized as a multidomain profile reflecting the presence of academic and uncertainty-related vulnerabilities together with the availability of psychological resources that may protect students when responding to academic and psychosocial demands. Accordingly, MIND-RISK is not intended to represent psychiatric diagnosis, but rather to identify patterns of vulnerability and protective capacity that may inform further assessment and counseling-related decision making.

Several established instruments have been widely used to assess psychological functioning among university students. The *Depression Anxiety Stress Scale* (DASS-21) assesses symptoms of depression, anxiety, and stress (Lovibond, 1995), whereas the *Patient Health Questionnaire-9* (PHQ-9) and *Generalized Anxiety Disorder-7* (GAD-7) were developed primarily to assess depressive and generalized anxiety symptoms, respectively (Kroenke et al., 2001; Spitzer et al., 2006). In contrast, the *Mental Health Continuum–Short Form* (MHC-SF) emphasizes positive mental health through emotional, social, and psychological well-being (Keyes, 2002), while the *Connor–Davidson Resilience Scale* (CD-RISC) focuses specifically on resilience as a capacity to cope with adversity (Connor & Davidson, 2003). These instruments have made important contributions to the assessment of particular dimensions of student mental health; however, they were developed to measure specific constructs or

domains rather than to function as an integrated academic-contextual risk index combining vulnerability factors and protective psychological resources within a single framework.

In the Indonesian higher education context, mental health screening has also been applied to university students to identify psychological difficulties and inform counseling-related support (Astuti et al., 2024). However, commonly used screening approaches tend to focus primarily on psychological symptoms or specific domains of functioning. Although such instruments are valuable for detecting particular forms of distress or psychological resources, they do not simultaneously integrate academic-contextual risk, uncertainty-related vulnerability, and multiple protective psychological resources within a single assessment framework. This limitation provides a contextual rationale for developing MIND-RISK as a complementary multidomain assessment tool rather than as a replacement for established symptom-based measures. Consequently, there remains a need for an assessment model that simultaneously considers academic risk, uncertainty-related vulnerability, adaptive psychological capacities, and positive functioning. MIND-RISK was developed to address this gap by integrating these dimensions into a single framework for identifying differences in university students' mental health risk and psychological resources. The development and empirical positioning of the MIND-RISK instrument are summarized in Table 1.

Table 1. State of the Art in Mental Health Assessment among University Students in Higher Education

Study	Research Focus	Key Findings	Research Gap	Contribution of the Present Study
Duraku et al., (2024)	Review of mental health services in higher education	Higher education institutions require early detection and assessment systems as part of university students' mental health services.	The study did not propose an assessment instrument integrating risk and protective factors within a single framework.	MIND-RISK integrates risk and protective dimensions into a single assessment instrument.
Abulfaraj et al., (2024)	Systematic review of resilience interventions among university students	Resilience interventions were effective in reducing stress and anxiety, although their effects on psychological well-being were inconsistent.	The focus was on intervention effectiveness rather than the development of an instrument for assessing mental health risk.	MIND-RISK assesses both mental health risk factors and protective psychological resources among university students.
Baik et al., (2026)	Systematic review of mental health assessment instruments for university students	Mental health assessment remains dominated by instruments that measure psychological constructs separately and are primarily oriented toward clinical conditions.	There is limited evidence of instruments integrating risk factors, protective factors, and academic context within a single measurement model.	MIND-RISK integrates the <i>Risk Factor Index</i> and <i>Protective Factor Index</i> within a multidomain assessment framework for university students.
Biscaldi et al., (2026)	Systematic review of instruments assessing well-being, ill-being, and health behaviors among university students	Most instruments assess individual dimensions of mental health and do not provide a comprehensive risk index.	A comprehensive index capable of representing university students' overall mental health risk remains limited.	MIND-RISK integrates the <i>Risk Factor Index</i> and <i>Protective Factor Index</i> within a multidomain assessment framework for university students..

Table 1 shows that research on the assessment of university students' mental health has expanded, but most studies continue to adopt a partial approach by assessing psychological constructs separately. To date, there remains limited evidence of instruments that simultaneously assess academic and psychological risk factors together with protective psychological resources within an integrated framework that can support decision-making in university counseling and mental health services.

Addressing this gap, MIND-RISK integrates six complementary constructs: *academic stress load* and *intolerance of uncertainty* as risk factors, and *academic adaptability*, *emotion regulation*, *psychological resilience*, and *psychological flourishing* as protective factors. The novelty of this study lies in integrating academic and psychological risk factors with protective psychological resources within a multidomain assessment framework for university

students. This approach captures students' levels of vulnerability while simultaneously considering adaptive psychological capacities that may contribute to their ability to manage academic and psychosocial demands. Accordingly, MIND-RISK is expected to serve as a basis for *decision-support assessment* in university counseling and mental health services.

This study aims to evaluate the psychometric characteristics of the MIND-RISK instrument, including item fit, measurement reliability, response-category functioning, and residual dependence, within its six-construct risk-protective framework for university students. The instrument is expected to demonstrate adequate psychometric quality while providing comprehensive information regarding students' risk characteristics and protective psychological resources. Academically, this study contributes to the advancement of mental health assessment in the field of counseling, while practically, MIND-RISK is expected to support early detection and evidence-based, contextually informed decision-making in university mental health services.

Method

This study employed a quantitative instrument-development design to evaluate the psychometric characteristics of the Mental Health INdex for Academic Distress and Psychological RISK (MIND-RISK) among university students. The development procedure included construct definition, indicator identification, item development, expert content validation, field testing, and psychometric evaluation using the Rasch model. Such a procedure is consistent with contemporary approaches to instrument development in which theoretical conceptualization and content representation are followed by empirical evaluation of item functioning and measurement quality (Hughes et al., 2022; Mohamat et al., 2022). Rasch analysis was selected because it provides item and person estimates on a common logit scale and enables examination of item fit, response-category functioning, measurement reliability, and residual dependence within a probabilistic measurement framework (Chong et al., 2022). Given the polytomous five-category response format, item parameters were estimated using the Partial Credit Model (PCM2).

The field test involved 100 university students from Universitas Islam Balitar and Universitas Sebelas Maret who were selected using purposive sampling. The inclusion criteria were: (1) being an active university student, (2) being at least 18 years old, (3) completing all 48 MIND-RISK items, and (4) providing voluntary consent to participate in the study. The sample consisted of 24 male students (24%) and 76 female students (76%), with participants ranging in age from 18 to 41 years. Prior to analysis, demographic responses with different capitalization, abbreviations, or wording but representing the same category were standardized to ensure consistency in participant characteristics. All participants received information regarding the study objectives, procedures, confidentiality of their responses, and their right to withdraw from participation at any time. The sample was used to provide empirical evidence regarding item functioning and the initial measurement properties of MIND-RISK rather than to establish population norms. The stability of item parameter estimates and measurement reliability in Rasch analysis may be influenced by sample size and the distribution of respondent measures (Wulandari & Hidayat, 2023). The original study design similarly positioned the field test as an initial psychometric evaluation rather than a normative study.

The sample size was positioned within the developmental purpose of the present study. The $N = 100$ calibration was intended to provide strengthened evidence of item functioning and measurement properties relative to the preliminary calibration, rather than to establish definitive population parameters or normative cut-off scores. In Rasch measurement, the stability of item estimates is influenced not only by sample size but also by test length, person-item targeting, and the distribution of respondent measures (Bintang & Suprananto, 2024). Therefore, the present sample was considered appropriate for the current instrument-development stage, while broader calibration with larger and more heterogeneous samples remains necessary before normative or standardized interpretation.

MIND-RISK was developed based on a theoretical framework integrating risk and protective factors associated with university students' mental health. The instrument comprises six constructs: Academic Stress Load, Intolerance of Uncertainty, Academic Adaptability, Emotion Regulation, Psychological Resilience, and Psychological Flourishing. These constructs were organized into two complementary dimensions. The Risk Factor Index comprises Academic Stress Load and Intolerance of Uncertainty, whereas the Protective Factor Index comprises Academic Adaptability, Emotion Regulation, Psychological Resilience, and Psychological Flourishing. Psychometric analyses were conducted at two complementary levels. First, the 16-item Risk Factor Index and the 32-item Protective Factor Index were calibrated separately. Second, the six theoretically specified constructs were analyzed independently: Academic Stress Load (I1–I8), Intolerance of Uncertainty (I9–I16), Academic Adaptability (I17–I24), Emotion Regulation (I25–I32), Psychological Resilience (I33–I40), and Psychological Flourishing (I41–I48). This construct-based strategy was used because favorable and unfavorable

wording represents item direction rather than the substantive latent construct being measured. The conceptualization of MIND-RISK followed the principle that psychological measurement should begin with a clearly defined construct domain and an explicit correspondence between theoretical indicators and item content (Hughes et al., 2022; Mohamat et al., 2022).

The initial version of MIND-RISK consisted of 48 items, comprising 24 favorable and 24 unfavorable items. The items were developed from indicators and behavioral descriptors derived from the operational definitions of the six constructs. Item scoring was adjusted according to the theoretical direction of each index. Within the Risk Factor Index, items were scored so that higher scores represented greater psychological risk, whereas within the Protective Factor Index, higher scores represented stronger protective psychological resources. Reverse scoring was applied where necessary to maintain the intended direction within each index. The instrument employed a five-point Likert-type response scale ranging from Strongly Inappropriate (1), Inappropriate (2), Uncertain (3), Appropriate (4), to Strongly Appropriate (5). Prior to Rasch estimation in TAM, the scored responses were recoded from the original 1–5 format to 0–4. This transformation changed only the numerical category labels and did not alter the ordering or substantive interpretation of the responses. Because ordered response categories require respondents to meaningfully distinguish between adjacent response options, the functioning of the rating scale was subsequently examined using Rasch category-functioning analysis (Chong et al., 2022).

Prior to field testing, the initial item pool was reviewed qualitatively by two experts, one in assessment and instrument development and one in guidance and counseling. The experts evaluated item relevance, wording clarity, correspondence with indicators, and content representativeness. All items were considered appropriate for field testing, although five items were revised based on the experts' recommendations before data collection. Because the content-validation stage employed qualitative expert judgment, no quantitative inter-rater agreement coefficient, CVR, or content-validity index was calculated. The overall development and psychometric evaluation process of MIND-RISK is summarized in Figure 1.

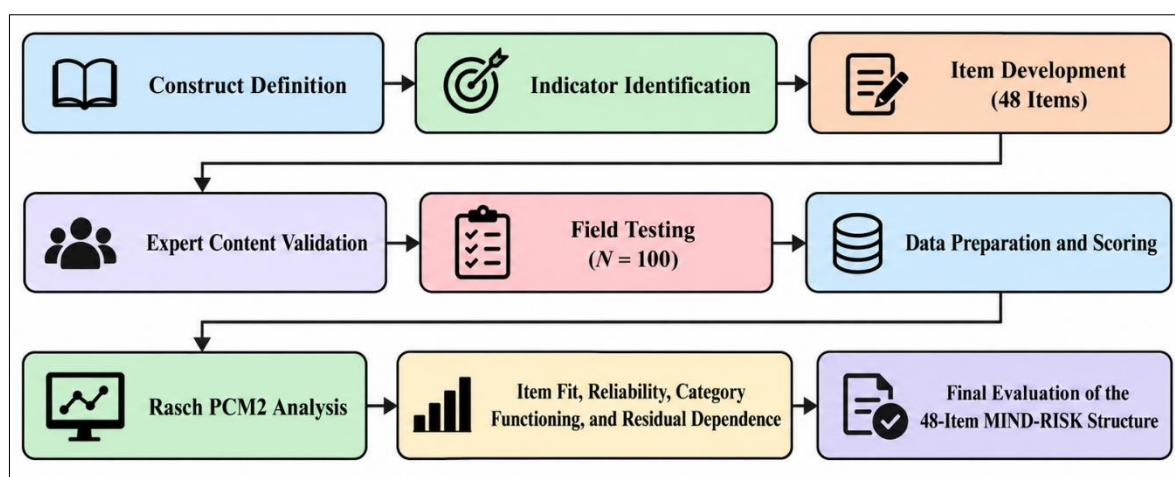


Figure 1. Development and Psychometric Evaluation Procedure of MIND-RISK

Field testing was conducted using an online questionnaire administered through Google Forms. Participants first completed demographic information and subsequently responded to all MIND-RISK items. The collected responses were examined for completeness, coded, and transferred into Microsoft Excel format prior to analysis. Demographic data were cleaned and standardized where differences in capitalization, abbreviations, or wording represented the same response category. Unfavorable items were reverse-scored according to their scoring direction. The cleaned dataset containing responses from 100 participants to all 48 items was analyzed in R version 4.6.1 using the TAM package version 4.3-25. Polytomous Rasch models were estimated using the PCM2 parameterization. Separate models were estimated for the Risk Factor Index, the Protective Factor Index, and each of the six theoretically specified constructs.

The Rasch model was used to examine the empirical functioning of individual items and the overall measurement properties of MIND-RISK. Rasch analysis enables item characteristics and respondent measures to be estimated on a common logit scale while providing diagnostic information regarding item fit, item location, response-category functioning, measurement reliability, and residual dependence (Wulandari & Hidayat, 2023). Previous instrument-development studies have also demonstrated the application of Rasch analysis for evaluating item validity, reliability, rating-scale functioning, and item hierarchy (Mulyanti et al., 2022).

Item quality was evaluated primarily using Infit and Outfit Mean Square statistics. Items were considered to demonstrate acceptable item-level fit when both Infit and Outfit MNSQ values fell within the predetermined range of 0.50–1.50. Item-location estimates in logits and their standard errors were also examined to describe the relative position of items within each independently calibrated construct. Infit MNSQ was used to evaluate unexpected responses that were more sensitive to observations near the respondent's estimated measure, whereas Outfit MNSQ was used to identify unexpected responses that were more sensitive to outlying observations. Pt-Measure Corr was used to examine whether each item contributed positively to the measurement of the intended attribute. Similar fit criteria have been applied in recent Rasch-based instrument studies when evaluating the empirical functioning of individual items (Asradi & Sarman, 2026).

Item evaluation was conducted iteratively by considering both psychometric evidence and substantive theoretical considerations. Items demonstrating potential psychometric problems were reviewed using multiple sources of evidence, including item fit, response-category functioning, reliability, residual dependence, and substantive theoretical relevance. Item deletion was considered only when converging psychometric evidence indicated inadequate functioning and when removal would not compromise representation of the intended construct. Decisions regarding item retention or elimination were not based exclusively on statistical indices but also considered the substantive contribution of each item to its intended construct and the representation of the six MIND-RISK constructs. Statistical item-fit indices were therefore interpreted together with theoretical and content considerations rather than being treated as the sole criteria for item retention or deletion (Hughes et al., 2022; Mohamat et al., 2022).

The functioning of the five-point response scale was examined using category-threshold analysis within the Partial Credit Model. Estimated category thresholds were examined to determine whether successive response categories operated in an ordered manner along the latent continuum. The analysis evaluated whether respondents were able to differentiate between successive response categories and whether the category structure corresponded to increasing levels of the measured attribute. Properly functioning rating categories should demonstrate an ordered progression of category thresholds, indicating that the response options provide meaningful distinctions along the latent continuum (Chong et al., 2022). Thresholds were obtained for each item using the TAM threshold estimation procedure. A monotonic sequence from Cat1 to Cat4 was interpreted as evidence of ordered category functioning. Evaluation of response-category functioning is particularly relevant for polytomous instruments because poorly functioning categories may reduce measurement precision and obscure distinctions between levels of the construct (Wulandari & Hidayat, 2023).

Measurement reliability was evaluated using Expected A Posteriori (EAP) reliability and Weighted Likelihood Estimate (WLE) reliability. EAP reliability was obtained from the marginal maximum-likelihood Rasch model and reflects the precision with which respondent locations were estimated from the model. WLE reliability was calculated from weighted-likelihood person estimates as complementary evidence of the reproducibility of individual respondent measures. Reliability estimates were examined separately for the Risk Factor Index, Protective Factor Index, and each of the six constructs.

Residual dependence and overall model fit were examined using residual-based diagnostics provided by TAM. Yen-type Q3 residual correlations and adjusted Q3 (aQ3) statistics were examined to identify residual associations between item pairs after accounting for the latent construct. The mean absolute adjusted Q3 (MADaQ3), maximum adjusted Q3, and standardized root mean square residual (SRMR) were also reported as complementary indicators of model fit and local dependence. These diagnostics were interpreted as evidence regarding residual dependence rather than as automatic criteria for item deletion.

Throughout the analysis, statistical quality and substantive construct representation were considered simultaneously. Item evaluation therefore considered both empirical fit and the extent to which each item contributed to the theoretical representation of the six MIND-RISK constructs and maintained appropriate coverage of the Risk Factor Index and Protective Factor Index. This approach recognizes that psychometric evaluation requires both empirical evidence and theoretical judgment so that decisions regarding item retention, revision, or removal do not compromise construct coverage (Hughes et al., 2022; Mohamat et al., 2022).

Results and Discussions

A total of 100 university students participated in the psychometric evaluation of the MIND-RISK instrument. The participants consisted of 24 male students (24.0%) and 76 female students (76.0%). All 100 response records were included in the Rasch analysis. The initial MIND-RISK instrument consisted of 48 items organized into six theoretically specified constructs. Rather than evaluating items according to favorable or unfavorable wording, the present analysis calibrated the items according to their substantive measurement structure: *Academic Stress Load* and *Intolerance of Uncertainty* as components of the Risk Factor Index, and *Academic*

Adaptability, Emotion Regulation, Psychological Resilience, and Psychological Flourishing as components of the Protective Factor Index.

Table 2. Respondent Characteristics by Gender

Gender	Frequency (n)	Percentage (%)
Male	24	24.0
Female	76	76.0
Total	100	100.0

The initial higher-order calibration was conducted separately for the 16-item Risk Factor Index and the 32-item Protective Factor Index using the Partial Credit Model. The Risk Factor Index yielded an EAP reliability of 0.835 and a WLE reliability of 0.833. All 100 respondents and all 16 risk items were included in the estimation. The Protective Factor Index demonstrated higher measurement precision, with an EAP reliability of 0.923 and a WLE reliability of 0.915.

Because both higher-order indices consist of theoretically distinguishable constructs, the analysis was subsequently conducted at the six-construct level. This allowed item functioning to be evaluated within the specific domain each item was intended to measure.

Table 3. Reliability of the MIND-RISK Indices and Constructs

Dimension / Construct	Items	Number of Items	EAP Reliability	WLE Reliability
Risk Factor Index	I1–I16	16	0.835	0.833
Academic Stress Load	I1–I8	8	0.800	0.799
Intolerance of Uncertainty	I9–I16	8	0.655	0.652
Protective Factor Index	I17–I48	32	0.923	0.915
Academic Adaptability	I17–I24	8	0.671	0.718
Emotion Regulation	I25–I32	8	0.671	0.713
Psychological Resilience	I33–I40	8	0.788	0.794
Psychological Flourishing	I41–I48	8	0.791	0.795

The WLE estimates confirmed that measurement precision was strongest for the higher-order Risk and Protective Factor Indices and somewhat lower when each eight-item construct was evaluated independently. Among the individual constructs, *Academic Stress Load*, *Psychological Resilience*, and *Psychological Flourishing* produced WLE reliability values close to 0.80. *Intolerance of Uncertainty* showed the lowest reliability (0.652), whereas *Academic Adaptability* and *Emotion Regulation* produced values of 0.718 and 0.713, respectively.

Item-level fit was evaluated using Outfit and Infit Mean Square statistics. The predetermined acceptable range was 0.50–1.50. All 16 items representing the Risk Factor Index met this criterion when calibrated within their respective constructs.

Within *Academic Stress Load*, Outfit MNSQ values ranged from 0.905 to 1.095 and Infit MNSQ values ranged from 0.905 to 1.076. Within *Intolerance of Uncertainty*, Outfit values ranged from 0.862 to 1.121 and Infit values ranged from 0.877 to 1.097. Thus, none of the 16 risk-factor items demonstrated item-level misfit.

Table 4. Rasch Item Measures and Fit Statistics for the Risk Factor Constructs

Construct	Item	Measure (logit)	SE	Outfit MNSQ	Infit MNSQ	Decision
Academic Stress Load	I1	0.785	0.141	1.083	1.042	Retained
	I2	-0.910	0.123	0.970	0.990	Retained
	I3	0.651	0.124	1.095	1.063	Retained
	I4	-0.270	0.110	1.077	1.076	Retained
	I5	0.789	0.130	0.905	0.905	Retained
	I6	0.042	0.115	1.008	1.023	Retained
	I7	0.473	0.131	1.064	0.955	Retained
	I8	-0.312	0.112	0.921	0.934	Retained
Intolerance of Uncertainty	I9	0.241	0.098	0.920	0.946	Retained
	I10	-0.533	0.097	0.936	0.941	Retained
	I11	0.044	0.104	1.121	1.097	Retained
	I12	-0.861	0.113	0.862	0.877	Retained
	I13	0.868	0.121	1.098	1.070	Retained
	I14	-0.297	0.108	1.068	1.070	Retained
	I15	0.068	0.109	1.056	1.063	Retained
	I16	-0.343	0.107	0.881	0.879	Retained

The item-location estimates showed variability within both risk constructs. For *Academic Stress Load*, the measures ranged from -0.910 logits for I2 to 0.789 logits for I5. Within *Intolerance of Uncertainty*, I12 had the lowest item measure (-0.861 logits), whereas I13 had the highest measure (0.868 logits). Item-level Outfit and Infit values for all 16 items remained within the specified fit interval.

Because the six constructs were estimated separately, these logit locations describe the relative position of items within each construct and should not be interpreted as directly comparable difficulty estimates across independently calibrated constructs.

All 32 protective-factor items also satisfied the item-level MNSQ criterion. For *Academic Adaptability*, Outfit MNSQ ranged from 0.933 to 1.260 and Infit MNSQ ranged from 0.939 to 1.229. Item I20 showed the highest values in this construct (Outfit = 1.260; Infit = 1.229), but both remained below the upper criterion of 1.50.

For *Emotion Regulation*, Outfit MNSQ ranged from 0.851 to 1.145 and Infit MNSQ ranged from 0.862 to 1.118. *Psychological Resilience* showed Outfit values between 0.838 and 1.283 and Infit values between 0.849 and 1.201. I35 exhibited the highest fit statistics within this construct but remained acceptable. Finally, *Psychological Flourishing* showed Outfit MNSQ values ranging from 0.843 to 1.192 and Infit values ranging from 0.860 to 1.188.

Table 5. Rasch Item Measures and Fit Statistics for the Protective Factor Constructs

Construct	Item	Measure (logit)	SE	Outfit MNSQ	Infit MNSQ	Decision
Academic Adaptability	I17	-0.466	0.123	0.933	0.963	Retained
	I18	0.259	0.112	1.034	0.991	Retained
	I19	-0.510	0.131	1.037	1.052	Retained
	I20	-0.144	0.110	1.260	1.229	Retained
	I21	-0.748	0.125	0.945	0.949	Retained
	I22	0.212	0.114	0.953	0.950	Retained
	I23	-0.719	0.137	1.025	1.013	Retained
	I24	-0.095	0.113	0.947	0.939	Retained
Emotion Regulation	I25	-0.479	0.128	1.145	1.113	Retained
	I26	0.164	0.110	1.139	1.118	Retained
	I27	-0.878	0.136	0.996	1.002	Retained
	I28	-0.176	0.113	1.089	1.094	Retained
	I29	-0.322	0.109	1.005	0.999	Retained
	I30	0.472	0.108	0.851	0.868	Retained
	I31	-0.269	0.105	0.920	0.935	Retained
	I32	0.517	0.110	0.852	0.862	Retained
Psychological Resilience	I33	-1.157	0.140	0.875	0.890	Retained
	I34	-0.399	0.117	0.838	0.849	Retained
	I35	-0.496	0.132	1.283	1.201	Retained
	I36	0.056	0.116	0.965	0.970	Retained
	I37	-1.298	0.142	1.011	0.992	Retained
	I38	0.166	0.115	1.181	1.111	Retained
	I39	-1.084	0.133	1.002	0.960	Retained
	I40	-0.348	0.117	1.051	1.047	Retained
Psychological Flourishing	I41	-1.072	0.123	1.019	1.053	Retained
	I42	-0.703	0.119	0.948	0.962	Retained
	I43	-0.732	0.122	1.030	1.053	Retained
	I44	-0.243	0.110	0.907	0.910	Retained
	I45	-0.444	0.133	1.079	1.077	Retained
	I46	0.103	0.106	1.192	1.188	Retained
	I47	-1.180	0.138	1.025	1.011	Retained
	I48	-0.009	0.111	0.843	0.860	Retained

The item measures for *Academic Adaptability* ranged from -0.748 to 0.259 logits, whereas *Emotion Regulation* ranged from -0.878 to 0.517 logits. The range for *Psychological Resilience* was -1.298 to 0.166 logits, and the corresponding range for *Psychological Flourishing* was -1.180 to 0.103 logits.

Across the six constructs, therefore, none of the 48 items exceeded the predefined item-level Outfit or Infit MNSQ range of 0.50–1.50. Consequently, no item was eliminated on the basis of item-level fit.

The five-point response structure was further examined using estimated category thresholds. For all 48 items, the thresholds obtained from `tam.threshold()` progressed in the expected order from Cat1 through Cat4. For example, the threshold sequence for I1 was -1.843, 0.376, 2.069, and 2.587 logits, while the sequence for I17

was -1.722, -1.415, -0.557, and 1.734 logits. Ordered threshold patterns were also observed throughout *Emotion Regulation*, *Psychological Resilience*, and *Psychological Flourishing*.

One category-specific fit deviation was observed for the third step of I3, which showed an Outfit MNSQ of 1.629. However, the overall I3 item fit remained acceptable (Outfit = 1.095; Infit = 1.063), and its estimated category thresholds remained ordered. Therefore, this step-level deviation did not result in item elimination at this stage.

Residual-based model-fit diagnostics revealed additional departures from strict local independence. The maximum adjusted Q3 ranged from 0.371 for *Intolerance of Uncertainty* to 0.530 for *Psychological Resilience*. SRMR values ranged from 0.112 to 0.169. The adjusted Q3 tests were statistically significant for all six construct-level models.

Table 6. Response Category Functioning and Residual Model-Fit Statistics

Construct	Cat1 Range	Cat2 Range	Cat3 Range	Cat4 Range	Threshold Order	MADaQ3	Max aQ3	p	SRMR
Academic Stress Load	-3.140 to -1.638	-1.379 to 0.376	-0.424 to 2.069	1.304 to 3.199	Ordered	0.175	0.430	<.001	0.112
Intolerance of Uncertainty	-2.263 to -1.356	-1.286 to 0.533	-0.633 to 1.355	0.730 to 2.939	Ordered	0.147	0.371	.003	0.126
Academic Adaptability	-3.152 to -1.722	-1.569 to -0.026	-0.758 to 0.867	1.665 to 2.396	Ordered	0.197	0.394	.001	0.152
Emotion Regulation	-2.519 to -1.488	-2.018 to 0.369	-0.459 to 0.859	1.433 to 3.055	Ordered	0.172	0.452	<.001	0.146
Psychological Resilience	-3.402 to -2.087	-2.165 to 0.190	-1.315 to 0.705	1.010 to 2.231	Ordered	0.248	0.530	<.001	0.169
Psychological Flourishing	-3.119 to -1.505	-2.148 to -0.425	-0.838 to 0.336	0.764 to 2.424	Ordered	0.222	0.511	<.001	0.144

The residual diagnostics for *Academic Stress Load* yielded a maximum adjusted Q3 of 0.4298 and an SRMR of 0.112, whereas *Intolerance of Uncertainty* yielded a maximum adjusted Q3 of 0.3713 and an SRMR of 0.126. *Academic Adaptability* and *Emotion Regulation* yielded maximum adjusted Q3 values of 0.3936 and 0.4518, respectively. The highest residual-dependence statistics were found for *Psychological Resilience* (max aQ3 = 0.5301) and *Psychological Flourishing* (max aQ3 = 0.5109).

At the higher-order index level, residual dependence was also present. The 16-item Risk Factor Index produced a maximum adjusted Q3 of 0.4702 and an SRMR of 0.136, whereas the 32-item Protective Factor Index produced a maximum adjusted Q3 of 0.5451 and an SRMR of 0.149. These results were retained as additional evidence regarding the multidomain measurement structure rather than being used as an automatic criterion for deleting otherwise well-fitting individual items.

Based on the N = 100 calibration, all 48 initial MIND-RISK items satisfied the item-level Outfit and Infit MNSQ criteria within their theoretically specified constructs. Furthermore, the estimated category thresholds remained ordered for all items. Accordingly, no item was removed during the present calibration, resulting in the retention of 16 Risk Factor items and 32 Protective Factor items.

Table 7. Composition of MIND-RISK Following the N = 100 Rasch Calibration

Higher-Order Index	Construct	Item Numbers	Favorable	Unfavorable	Retained Items
Risk Factor Index	Academic Stress Load	I1–I8	4	4	8
	Intolerance of Uncertainty	I9–I16	4	4	8
Risk Factor subtotal		I1–I16	8	8	16
Protective Factor Index	Academic Adaptability	I17–I24	4	4	8
	Emotion Regulation	I25–I32	4	4	8
	Psychological Resilience	I33–I40	4	4	8
	Psychological Flourishing	I41–I48	4	4	8
Protective Factor subtotal		I17–I48	16	16	32
Total MIND-RISK	6 constructs	I1–I48	24	24	48

Thus, the N = 100 analysis resulted in 48 retained items and zero eliminated items, replacing the previous preliminary 42-item solution obtained from the smaller calibration sample. Importantly, item retention was based on item functioning within the six theoretically specified constructs rather than on whether an item was

favorably or unfavorably worded. The residual-dependence findings were documented for further interpretation but did not, by themselves, provide sufficient item-level evidence for deletion.

The present findings provide updated psychometric evidence for the MIND-RISK instrument based on a larger calibration sample of 100 university students. In contrast with the preliminary analysis, which resulted in the elimination of several items, the current construct-based Rasch analysis supported the retention of all 48 items. All items demonstrated acceptable Outfit and Infit MNSQ values within the predetermined range of 0.50–1.50 when evaluated within their theoretically specified constructs. This finding indicates that the original item pool remained empirically compatible with the measurement model when the six theoretical domains of MIND-RISK were taken into account. The result is consistent with recent applications of Rasch analysis in instrument development, which emphasize the use of item-fit information to evaluate item functioning while maintaining the substantive structure of the construct being measured (Rozenal et al., 2023).

To position the present findings within the broader psychometric literature, Table 8 provides a structured comparison with 15 validation studies involving psychological and mental-health-related instruments, particularly those developed or evaluated among university students. The comparison includes Rasch-based studies as well as validation studies examining external validity and measurement invariance, thereby providing a broader context for interpreting the current psychometric evidence for MIND-RISK.

Table 8. Comparison of Psychometric Validation Studies Relevant to MIND-RISK

Study	Instrument / Construct	Sample and Approach	Main Psychometric Finding
Asradi & Sarman (2026)	Academic Resilience Scale	606 Indonesian university students; Rasch	The 26-item scale showed good person and item reliability, acceptable item fit, unidimensionality, and no meaningful gender DIF.
Oktasari et al., (2026)	Indonesian Resilience at University Scale (ID-RAUS)	1,046 Indonesian university students from 19 provinces; Rasch	The adapted scale demonstrated satisfactory validity and reliability, supporting culturally contextualized resilience measurement in Indonesian higher education.
Nawantara et al., (2026)	Ryff Psychological Well-Being Scale	1,215 Indonesian university students; Rasch	Three of the original 42 items showed persistent misfit and were removed, resulting in a 39-item version with satisfactory psychometric properties.
Purnama et al., (2022)	Ryff Psychological Well-Being Scale	425 Indonesian higher education students; Rasch	Person reliability was 0.85 and item reliability 0.99, although the authors identified limitations in person separation and factorial representation.
Liu & Lim (2022)	Brief Resilience Scale	839 and 1,068 Singaporean undergraduate freshmen; Rasch	Rasch analyses supported the scale's validity and reliability, while also indicating that additional difficult items could improve targeting at higher resilience levels.
Rozenal et al., (2023)	Perceived Stress Scale	793 university students; Rasch	Rasch analysis indicated dimensional differences between positively and negatively worded items, illustrating the importance of examining wording and dimensional structure rather than relying only on global reliability.
Ratnaningsih et al., (2025)	Psychological Capital Questionnaire	1,012 Indonesian undergraduate students; Rasch	The Indonesian PCQ demonstrated satisfactory item quality, person reliability, unidimensionality, and appropriate response-category functioning across its dimensions.
Yusuf et al., (2025)	Student Mental Health Scale	266 Indonesian university students; Rasch	The culturally contextualized scale was refined through expert review and Rasch evaluation and provided evidence for measuring student mental-health status in the Indonesian context.
Cassiani-Miranda et al., (2025)	PHQ-9	550 Colombian university students; Rasch	The PHQ-9 showed acceptable internal consistency, unidimensionality, local independence, and generally limited gender-related item bias.
White & Karr, (2025)	GAD-7	582 college students; factor analysis and measurement invariance	The GAD-7 showed excellent reliability, convergent and discriminant validity, and measurement invariance across several demographic groups.
Aruta et al., (2024)	Filipino DASS-21	681 university students; structural validation	The Filipino version retained the three-factor structure and demonstrated convergent, criterion-related, and item-level validity evidence.

Study	Instrument / Construct	Sample and Approach	Main Psychometric Finding
Gustin et al., (2025)	Hyper-Independence Scale	200 Indonesian university students; Rasch	The adaptation employed translation, expert review, pilot testing, and Rasch analysis to evaluate reliability, unidimensionality, and item functioning.
Sumintono et al., (2025)	Self-Handicapping Scale	318 Malaysian and 470 Indonesian university students; Rasch Rating Scale Model	The scale demonstrated satisfactory unidimensionality and measurement functioning across both national samples, including examination of rating-scale functioning and item bias.
Nadhirah et al., (2026)	PERMA-Based Flourishing Scale	305 university students; Rasch	Rasch analysis supported the validity and reliability of a multidimensional positive-functioning construct designed specifically for university students.
Mauladi et al., (2026)	Flourishing Scale	236 university students; Rasch	The scale showed person reliability of 0.89, item reliability of 0.99, and adequate unidimensionality, although three items required further revision because of misfit.

The comparative evidence in Table 8 shows that contemporary psychometric validation studies rarely rely on a single reliability or item-fit coefficient. Rasch-based studies among university students have commonly examined item fit, reliability, dimensionality, response-category functioning, targeting, and Differential Item Functioning as complementary sources of measurement evidence. Item deletion has also varied across studies. Nawantara et al., (2026) removed three Psychological Well-Being Scale items because of persistent misfit, whereas Mauladi et al., (2026) identified several flourishing items requiring revision. In contrast, Asradi & Sarman (2026) retained their academic resilience items when item-fit and dimensionality criteria were satisfied. These findings support the approach adopted in the present study, in which item retention was determined by converging psychometric and theoretical evidence rather than by a single statistical criterion.

A second distinction concerns the scope of the constructs being assessed. Existing instruments typically focus on a particular domain, such as resilience (Asradi & Sarman, 2026; Liu & Lim, 2022; Oktasari et al., 2026), psychological well-being or flourishing (Mauladi et al., 2026; Nadhirah et al., 2026; Nawantara et al., 2026), psychological capital (Ratnaningsih et al., 2025), self-handicapping (Sumintono et al., 2025), depression (Cassiani-Miranda et al., 2025), anxiety (White & Karr, 2025), or depression, anxiety, and stress symptoms (Aruta et al., 2024). MIND-RISK differs from these instruments by simultaneously representing two vulnerability constructs and four protective constructs within a single multidomain risk-protective framework. Thus, its contribution does not lie merely in providing another measure of psychological symptoms or positive functioning, but in allowing vulnerability and psychological resources to be interpreted together for preliminary assessment and counseling-related decision support.

At the same time, comparison with these validation studies highlights areas in which further evidence for MIND-RISK remains necessary. For example, White and Karr (2023) examined convergent, discriminant, and measurement-invariance evidence for the GAD-7, while Aruta et al., (2024) reported convergent and criterion-related evidence for the Filipino DASS-21. The present MIND-RISK study primarily provides internal measurement evidence from Rasch analysis and has not yet examined correspondence with established external measures. Therefore, the current findings should be interpreted as developmental psychometric evidence, with convergent, discriminant, criterion-related, predictive, and invariance analyses remaining priorities for subsequent validation.

An important methodological implication of the current findings is that MIND-RISK should not be evaluated primarily by separating favorable and unfavorable statements. The 48 items were originally balanced between 24 favorable and 24 unfavorable statements, and unfavorable items were reverse-scored before analysis to ensure the intended scoring direction. However, favorable or unfavorable wording represents item format rather than the substantive dimension being measured. In the present analysis, item functioning was therefore examined according to the six theoretically specified constructs: *Academic Stress Load*, *Intolerance of Uncertainty*, *Academic Adaptability*, *Emotion Regulation*, *Psychological Resilience*, and *Psychological Flourishing*. This construct-based approach produced a more theoretically coherent evaluation of item quality than grouping items solely on the basis of statement direction.

The retention of all 48 items should nevertheless not be interpreted as indicating that every item functioned identically. Some items showed relatively higher MNSQ values within their respective constructs. For example, I20 in *Academic Adaptability* showed an Outfit MNSQ of 1.260 and an Infit MNSQ of 1.229, I35 in *Psychological Resilience* showed an Outfit MNSQ of 1.283 and an Infit MNSQ of 1.201, and I46 in *Psychological Flourishing*

showed an Outfit MNSQ of 1.192 and an Infit MNSQ of 1.188. Nevertheless, these values remained within the specified acceptable range. Consequently, item deletion was not considered justified solely on the basis of relatively higher fit statistics. Recent psychometric research similarly emphasizes that item evaluation should consider several sources of measurement evidence rather than relying mechanically on a single fit statistic, including reliability, response-category functioning, dimensionality, targeting, and item bias (Sumintono et al., 2025).

The findings also provide an important perspective on the use of negatively worded items. All 24 unfavorable items remained within acceptable item-fit criteria after reverse scoring. This indicates that negative wording did not automatically produce psychometrically unacceptable items in the present sample. Therefore, unfavorable statements should not be considered inferior merely because of their linguistic direction. Instead, their quality should be judged according to their empirical fit and substantive contribution to the construct being measured. At the same time, negative wording can introduce additional cognitive response processes, particularly when respondents must interpret negations or distinguish statements that differ in direction. Thus, although the present results support the retention of the unfavorable items, their functioning should continue to be examined in subsequent validation studies.

The reliability results further demonstrate that the precision of MIND-RISK varied according to the level at which the instrument was evaluated. The 16-item Risk Factor Index produced a WLE reliability of 0.833, whereas the 32-item Protective Factor Index produced a higher WLE reliability of 0.915. These findings indicate relatively strong measurement precision when the constructs were represented within the two broader risk and protective indices. At the construct level, WLE reliability values were 0.799 for *Academic Stress Load*, 0.652 for *Intolerance of Uncertainty*, 0.718 for *Academic Adaptability*, 0.713 for *Emotion Regulation*, 0.794 for *Psychological Resilience*, and 0.795 for *Psychological Flourishing*.

The comparatively lower reliability of *Intolerance of Uncertainty* deserves particular attention. Its WLE reliability of 0.652 was lower than those of the other five constructs. However, the item-fit results for I9–I16 were consistently acceptable, with Outfit MNSQ values ranging from 0.862 to 1.121 and Infit MNSQ values ranging from 0.877 to 1.097. Therefore, the lower reliability does not appear to be attributable to an obviously misfitting individual item. Rather, it may indicate more limited differentiation among respondents within this eight-item domain or restricted variability in the latent attribute within the current sample. Accordingly, deleting an item merely to increase the reliability coefficient would not be warranted at this stage.

This finding is relevant in student mental health assessment because psychological vulnerability and psychological resources are not entirely dichotomous conditions. Students may occupy different positions across risk and protective dimensions, and variation may be more pronounced in some psychological domains than in others. Consequently, the reliability of an individual eight-item construct should be interpreted within the broader multidomain structure of MIND-RISK rather than as an isolated indicator of overall instrument quality. This perspective is consistent with evidence showing that university student mental health is associated with both stress-related vulnerabilities and protective resources such as resilience and social support (Fang et al., 2025; Turner & Holdsworth, 2024).

The distribution of item measures also demonstrated that the items occupied different locations within their respective construct continua. For example, item measures ranged from -0.910 to 0.789 logits in *Academic Stress Load*, from -0.861 to 0.868 logits in *Intolerance of Uncertainty*, from -0.748 to 0.259 logits in *Academic Adaptability*, and from -0.878 to 0.517 logits in *Emotion Regulation*. Similarly, *Psychological Resilience* ranged from -1.298 to 0.166 logits and *Psychological Flourishing* from -1.180 to 0.103 logits. These variations suggest that the retained items were not concentrated at identical locations within each latent continuum. Nevertheless, because each construct was calibrated independently, item-location estimates should primarily be interpreted within their respective domains rather than compared directly across different constructs. Measurement characteristics such as item hierarchy are also influenced by sample characteristics and person–item targeting and therefore require replication in broader samples (Bintang & Suprananto, 2024).

The functioning of the five-point response format provides another important improvement over the preliminary findings. In the present analysis, estimated category thresholds progressed in the expected order across all 48 items. This suggests that the five response categories generally represented progressively different levels of endorsement for the current sample. Ordered thresholds are important in polytomous Rasch analysis because they indicate that respondents are able to distinguish adjacent categories in an interpretable sequence. Previous psychometric research has emphasized that disordered thresholds may indicate difficulty distinguishing adjacent response categories and may require revision or collapsing of categories (Colledani et al., 2025). In the present study, however, the ordered thresholds provide support for retaining the existing five-category response structure.

A more detailed examination nevertheless identified an isolated category-level fit deviation. The third step of I3 showed an Outfit MNSQ of 1.629, even though the overall item fit of I3 remained acceptable and its category thresholds were ordered. This finding illustrates why item-retention decisions should not be based mechanically on a single category-level statistic. Category-level diagnostics are useful for identifying specific response transitions that may warrant attention, but they should be interpreted together with overall item functioning and category ordering. In this case, the broader evidence did not justify deletion of I3.

Despite the satisfactory item-level fit and ordered thresholds, the residual-based model diagnostics indicate that the measurement structure is not completely free from local dependence. Maximum adjusted Q3 values ranged from 0.371 for *Intolerance of Uncertainty* to 0.530 for *Psychological Resilience*, while SRMR values ranged from 0.112 to 0.169 across the six constructs. Relatively elevated residual associations were particularly evident for *Psychological Resilience* and *Psychological Flourishing*. These results suggest that certain pairs of items may share variance beyond the latent construct represented by the Rasch model.

The residual-dependence findings should not automatically be interpreted as evidence that particular items are invalid. Local dependence may arise because items address closely related experiences, use similar wording, or represent highly related facets within a broader psychological domain. More importantly, the item-level Infit and Outfit statistics remained within the acceptable range across all six constructs. The appropriate interpretation is therefore that individual item functioning was satisfactory, while the internal dependency structure requires additional investigation. This distinction is consistent with contemporary Rasch approaches that treat item fit, reliability, rating-scale functioning, dimensionality, targeting, and other diagnostic information as complementary rather than interchangeable sources of evidence (Rozental et al., 2023; Sumintono et al., 2025).

The same issue was observed when the constructs were combined into the two higher-order indices. The Risk Factor Index produced a maximum adjusted Q3 of 0.470, whereas the Protective Factor Index produced a maximum adjusted Q3 of 0.545. These findings are particularly important for interpreting the conceptual structure of MIND-RISK. Although the two indices provide useful organizing frameworks, they consist of theoretically distinct domains. The Risk Factor Index comprises *Academic Stress Load* and *Intolerance of Uncertainty*, while the Protective Factor Index comprises *Academic Adaptability*, *Emotion Regulation*, *Psychological Resilience*, and *Psychological Flourishing*. Thus, the higher-order indices should not be interpreted as evidence that all items within each index necessarily represent a strictly homogeneous unidimensional construct.

This multidomain structure is consistent with the conceptual objective of MIND-RISK. The instrument was designed not only to identify psychological vulnerability but also to assess psychological resources that may help students manage academic demands. The Risk Factor Index captures academic pressure and difficulty responding to uncertainty, whereas the Protective Factor Index captures adaptability, emotional regulation, resilience, and flourishing. Such an approach provides a broader perspective than screening frameworks based exclusively on psychological difficulties. Recent research among university students similarly indicates that mental well-being reflects the interplay between stressful experiences and protective resources, while resilience represents an important psychological resource under challenging circumstances (Abulfaraj et al., 2024; Turner & Holdsworth, 2024). Associations among emotion regulation, resilience, and positive psychological outcomes further support the inclusion of adaptive psychological resources within a broader assessment framework (Chen et al., 2022).

Accordingly, MIND-RISK can be conceptually positioned as both a risk-oriented screening instrument and a framework for identifying psychological resources. This orientation is particularly relevant to guidance and counseling services in higher education because assessment results can be used to identify students who may require attention while simultaneously indicating psychological capacities that could be strengthened through developmental and preventive services. Current developments in higher education mental health similarly emphasize approaches that combine identification of psychological difficulties with strengthening protective resources and improving accessible support systems (Pointon-Haas et al., 2024).

The findings therefore have practical implications for the use of MIND-RISK in university guidance and counseling settings. An instrument with acceptable item-level functioning can provide structured information regarding academic pressure, uncertainty, adaptive functioning, emotional regulation, resilience, and psychological flourishing. Such information may assist counselors in identifying areas that require additional assessment and in determining the focus of preventive, developmental, or responsive services. Within counseling practice, assessment-based decision making should remain grounded in professional and ethical considerations so that psychological information is used to support, rather than mechanically determine, counseling decisions (Kurniawan et al., 2026). Mental health screening in Indonesian higher education has similarly been used to identify psychological difficulties and inform counseling and stress-management support (Astuti et al., 2024).

However, MIND-RISK should not be positioned as a clinical diagnostic instrument. Its primary purpose at the present stage is preliminary assessment and mapping of students' risk characteristics and protective psychological resources. A high Risk Factor Index does not constitute evidence of a psychiatric disorder, just as a low Protective Factor Index should not be interpreted as a clinical diagnosis. Screening information should therefore be considered alongside interviews, observations, academic history, and other relevant psychological information when further assessment is necessary. This distinction is particularly important in higher education because the purpose of screening is early identification and appropriate support rather than replacement of professional diagnostic assessment.

From a service perspective, the six-construct structure provides a basis for more specific interpretation. Elevated *Academic Stress Load* may direct attention toward workload, time pressure, and coping with academic demands, whereas elevated *Intolerance of Uncertainty* may indicate a need to examine how students respond to unpredictable academic situations. On the protective side, lower *Academic Adaptability* may suggest difficulty adjusting to changing academic requirements, while lower *Emotion Regulation* or *Psychological Resilience* may point to psychological capacities that could be strengthened through counseling interventions or developmental services. Lower *Psychological Flourishing* may additionally indicate a need to explore students' sense of meaning, personal growth, satisfaction, and positive functioning. This type of domain-specific information is consistent with contemporary higher education mental health approaches that emphasize targeted and preventive support rather than relying exclusively on treatment after problems have become severe (Zhai & Carney, 2024).

Theoretically, the current findings strengthen the rationale for examining student mental health through the simultaneous consideration of vulnerability and psychological resources. All six constructs remained represented because all 48 original items were retained in the current $N = 100$ calibration. Thus, unlike the preliminary solution, preservation of conceptual coverage did not require accepting the loss of theoretically relevant items. The current structure provides a more complete empirical basis for examining risk and protection as related but distinguishable components of student mental health.

Psychometrically, the study also illustrates the value of evaluating MIND-RISK as a multidomain instrument rather than relying solely on a single global score or reliability coefficient. Rasch analysis provided information regarding individual item functioning, item locations, measurement precision, response-category thresholds, and residual dependence. Taken together, these findings indicate that evidence of measurement quality cannot be reduced to one statistic. Contemporary Rasch-based research similarly emphasizes the importance of considering item quality, reliability, rating-scale functioning, dimensionality, targeting, and item bias as complementary sources of evidence (Rozenal et al., 2023; Sumintono et al., 2025).

Several limitations nevertheless remain. First, although the current study increased the calibration sample from the preliminary stage to 100 students, this sample remains relatively limited for establishing stable population-level calibration. The distribution of participants was also unequal by gender, with substantially more female than male respondents. Consequently, the present findings should still be regarded as evidence from an instrument-development sample rather than definitive evidence of generalizability across university student populations. Sample size and sample characteristics can affect the stability of psychometric estimates, and larger heterogeneous samples remain important for stronger conclusions about item functioning (Bintang & Suprananto, 2024; Sumintono et al., 2025).

Second, the residual-dependence findings require more detailed examination. Although maximum adjusted Q3 identified residual associations within each domain, the present stage of analysis did not establish whether these associations result from overlapping content, similar wording, shared subfacets, or another source of local dependence. Future validation should therefore identify specific item pairs with elevated residual correlations and determine whether item revision, testlet modeling, or a more explicitly multidimensional measurement model provides a better representation of the data.

Third, additional evidence is required regarding Differential Item Functioning across relevant student characteristics. The present sample contained 24 male and 76 female participants, making gender-based subgroup comparisons possible only with caution because of the unequal group sizes. Broader validation should include more balanced and diverse participants across gender, study programs, educational levels, and higher education institutions. Further analyses should examine DIF, item targeting, dimensionality, and the stability of response-category functioning. Recent Rasch research involving university students similarly demonstrates the importance of investigating not only item fit and reliability but also dimensionality, targeting, and item bias when establishing stronger measurement evidence (Sumintono et al., 2025).

Fourth, the present study focused primarily on internal measurement evidence derived from Rasch analysis. Convergent, discriminant, criterion-related, and predictive validity were not examined using external psychological measures. Consequently, the present findings indicate how MIND-RISK items function within

the proposed measurement framework but do not yet establish the extent to which MIND-RISK scores correspond with established measures of psychological distress, well-being, or clinically relevant outcomes. Future validation should therefore examine relationships between MIND-RISK and established instruments such as measures of depression, anxiety, stress, psychological well-being, or related constructs.

Fifth, although retention of the complete 48-item structure preserved representation of all six theoretical constructs, the practical burden associated with completing the instrument was not formally evaluated in the present study. Completion time, respondent fatigue, and perceived ease of use should therefore be examined in subsequent studies. Future research may also investigate whether a shorter version of MIND-RISK can be developed without substantially reducing construct coverage or measurement precision.

Therefore, the present findings should be interpreted as strengthened but still developmental psychometric evidence for MIND-RISK. The $N = 100$ calibration supports retention of the 48-item structure, maintenance of the five-point response format, and preservation of the six-construct risk-protective framework. Nevertheless, the relatively lower reliability of Intolerance of Uncertainty, residual dependence, absence of external validity evidence, and unexamined respondent burden indicate areas requiring further investigation. Subsequent validation should involve larger and more heterogeneous samples, external criterion measures, detailed residual analysis, DIF testing, multidimensional model evaluation, and assessment of the practical feasibility of the 48-item format before MIND-RISK is positioned as a standardized screening instrument for broader use.

Conclusions

This study provides strengthened developmental psychometric evidence for MIND-RISK as a multidomain framework for assessing psychological vulnerability and protective resources among university students. Construct-based Rasch analysis supported the retention of all 48 items across six theoretically specified domains: Academic Stress Load, Intolerance of Uncertainty, Academic Adaptability, Emotion Regulation, Psychological Resilience, and Psychological Flourishing. The five-point response structure also demonstrated generally appropriate functioning, although differences in construct-level reliability and residual dependence indicate that further psychometric examination remains necessary. The main contribution of MIND-RISK lies in the simultaneous representation of academic and psychological vulnerability together with adaptive psychological resources within a single assessment framework. In contrast to instruments that primarily assess specific symptoms or individual psychological constructs, MIND-RISK is designed to provide a multidomain profile in which risk and protective characteristics can be considered together. This framework may support preliminary assessment and psychological profile mapping in university guidance and counseling services by helping counselors identify domains that require further assessment or developmental support. The present findings should not be interpreted as evidence that MIND-RISK is ready for clinical or standardized diagnostic use. Population norms, clinical cut-off scores, convergent and discriminant validity, criterion-related and predictive validity, and measurement invariance have not yet been established. Future studies should therefore involve larger and more heterogeneous samples, external psychological measures, Differential Item Functioning, multidimensional model evaluation, and examination of the practical feasibility of the 48-item format. These steps are necessary before MIND-RISK can be considered for broader standardized application across university populations.

Acknowledgments

The authors would like to express their sincere gratitude to the university students who voluntarily participated in this study and contributed their responses to the psychometric evaluation of the MIND-RISK instrument. The authors also gratefully acknowledge the experts who reviewed the content of the instrument and provided constructive feedback on the relevance, clarity, and representativeness of the items. The authors further acknowledge the individuals who provided technical assistance during data collection and data preparation for the Rasch analysis. All individuals acknowledged in this section have been informed of their inclusion and have given permission to be acknowledged. This research was funded by the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia (Kementerian Riset, Teknologi, dan Pendidikan Tinggi/Kemristekdikti) through the *Penelitian Dosen Pemula* (PDP) Research Grant Program in 2026, under Master Contract No. 286/C3/DT.05.00/PL-BARU/2026 and Derivative Contract No. 055/LL7/DT.05.00/PL-BARU/2026 and 097.06/KP.16/UNISBA/IV/2026, dated 13 April 2026 and 14 April 2026, respectively. The funding agency had no role in the study design, data collection, data analysis, interpretation of the results, or writing of the manuscript.

References

- Abulfaraj, G. G., Upsher, R., Zavos, H. M. S., & Dommett, E. J. (2024). The impact of resilience interventions on university students' mental health and well-being: A systematic review. *Education Sciences*, 14(5), 510. <https://doi.org/10.3390/educsci14050510>
- Ahmed, I., Hazell, C. M., Edwards, B., Glazebrook, C., & Davies, E. B. (2023). A systematic review and meta-analysis of studies exploring prevalence of non-specific anxiety in undergraduate university students. *BMC Psychiatry*, 23(1), 240. <https://doi.org/10.1186/s12888-023-04645-8>
- Akhzalini, H. A., Pambudi, P. R., & Risdiantoro, R. (2026). Makna Religiusitas sebagai Sistem Pembentukan Makna pada Mahasiswa yang Mengalami Quarter Life Crisis dalam Perspektif Psikologi. *Al-Isyrof: Jurnal Bimbingan Konseling Islam*, 8(1), 103–116. <https://doi.org/10.51339/isyrof.v8i1.4817>
- Aruta, J. J. B. R., Mateo, N. J., Resurreccion, R. R., Galanza, M. A. M. C., & Bernardo, A. B. I. (2024). Screening psychological symptoms in Filipino university students during the COVID-19 pandemic: Translation and structural validation of the Filipino version of the DASS-21. *Psychology in the Schools*, 61(8), 3243–3262. <https://doi.org/10.1002/pits.23214>
- Asradi, A., & Sarman, F. (2026). Psychometric Validation and Rasch Model Analysis of the Academic Resilience Scale for Indonesian University Students. *G-Couns: Jurnal Bimbingan Dan Konseling*, 10(03), 2406–2419. <https://doi.org/10.31316/g-couns.v10i03.8722>
- Astuti, F. D., Heriyanto, M. J., Desvita, W. R., Rokhmayanti, R., Hastuti, S. K. W., Utami, B. B. K., & Azka, A. (2024). Mental health screening for university students in the Special Region of Yogyakarta. *Journal of Epidemiology and Public Health*, 9(3), 343–353. <https://doi.org/10.26911/jepublichealth.2024.09.03.08>
- Baik, C., Ryan, T., & Lynch-Wells, S. (2026). Assessing student mental health and wellbeing in higher education: A systematic review of measurement approaches. *Educational Research Review*, 100823. <https://doi.org/10.1016/j.edurev.2026.100823>
- Bintang, R. S., & Suprananto, S. (2024). The impact of sample size, test length, and person-item targeting on the separation reliability in Rasch model: A simulation study. *JP3I (Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia)*, 13(1), 1–13. <https://doi.org/10.15408/jp3i.v13i1.27975>
- Biscaldi, V., Guerini, J., Ghelfi, M., & Velasco, V. (2026). Instruments used to measure well-being, ill-being, and health-related lifestyle behaviors in students attending Italian universities: a systematic review. *Frontiers in Public Health*, 14, 1787567. <https://doi.org/10.3389/fpubh.2026.1787567>
- Cassiani-Miranda, C. A., Scoppetta, O., Barrios-Villadiego, M. A., Tirado-Otálvaro, A. F., & Duran-Bedoya, A. C. (2025). Construct validity of the PHQ-9 in university students in Colombia: A Rasch analysis approach. *Medicina Clínica y Social*, 9(1), e521. <https://doi.org/10.52379/mcs.v9.521>
- Charbonnier, E., Montalescot, L., Puechlong, C., Goncalves, A., & Le Vigouroux, S. (2023). Relationship between Fear of COVID-19, Intolerance of Uncertainty, and Coping Strategies on University Students' Mental Health. In *Nutrients* (Vol. 15, Number 23, p. 4938). <https://doi.org/10.3390/nu15234938>
- Chen, X., Huang, Z., & Lin, W. (2022). The effect of emotion regulation on happiness and resilience of university students: The chain mediating role of learning motivation and target positioning. *Frontiers in Psychology*, Volume 13. <https://doi.org/10.3389/fpsyg.2022.1029655>
- Chong, J., Mokshein, S. E., & Mustapha, R. (2022). Applying the Rasch Rating Scale Model (RSM) to investigate the rating scales function in survey research instrument. *Cakrawala Pendidikan: Jurnal Ilmiah Pendidikan*, 41(1), 97–111. <https://doi.org/10.21831/cp.v41i1.39130>
- Colledani, Daiana, González Pizzio, Adriana P, Devita, Maria, & Anselmi, Pasquale. (2025). Investigating the Functioning of Rating Scales With Rasch Models. *Assessment*, 32(3), 434–446. <https://doi.org/10.1177/10731911241245792>
- Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, 18(2), 76–82. <https://doi.org/10.1002/da.10113>
- Duraku, Z. H., Davis, H., Arënliu, A., Uka, F., & Behluli, V. (2024). Overcoming mental health challenges in higher education: a narrative review. *Frontiers in Psychology*, 15, 1466060. <https://doi.org/10.3389/fpsyg.2024.1466060>
- Fang, S., Barker, E., Arasaratnam, G., Lane, V., Rabinovich, D., Panaccio, A., O'Connor, R. M., Nguyen, C. T., & Doucerain, M. M. (2025). Resilience, Stress, and Mental Health Among University Students: A Test of the Resilience Portfolio Model. *Stress and Health*, 41(2), e3508. <https://doi.org/10.1002/smi.3508>
- Gustin, V. Y., Nurwidawati, D., & Agustin, N. N. (2025). Adaptation and validation of the hyper-independence scale among Indonesian university students using the Rasch model. *REID (Research and Evaluation in Education)*, 11(2), 170–182. <https://doi.org/10.21831/reid.v11i2.90977>
- Huang, A., Liu, L., Wang, X., Chen, J., Liang, S., Peng, X., Li, J., Luo, C., Fan, F., & Zhao, J. (2023). Intolerance of uncertainty and anxiety among college students during the re-emergence of COVID-19: Mediation effects of cognitive emotion regulation and moderation effects of family function. *Journal of*

- Affective Disorders*, 327, 378–384. <https://doi.org/10.1016/j.jad.2023.01.110>
- Hughes, S. E., Haroon, S., Subramanian, A., McMullan, C., Aiyegbusi, O. L., Turner, G. M., Jackson, L., Davies, E. H., Frost, C., McNamara, G., Price, G., Matthews, K., Camaradou, J., Ormerod, J., Walker, A., & Calvert, M. J. (2022). Development and validation of the symptom burden questionnaire for long covid (SBQ-LC): Rasch analysis. *BMJ*, 377, e070230. <https://doi.org/10.1136/bmj-2022-070230>
- Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Social Behavior*, 43(2), 207–222. <https://doi.org/10.2307/3090197>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kurniawan, N. A., Rianti, D. D., Hanafi, H., Widyatmoko, W., & Pambudi, P. R. (2026). Pendidikan Kedamaian sebagai Fondasi Konseling Karier: Studi Literatur tentang Etika Konselor dan Pengambilan Keputusan Siswa. *Jurnal Bimbingan Dan Konseling Pandohop*, 6(1), 45–59. <https://doi.org/10.37304/pandohop.v6i1.23075>
- Leow, T., Li, W. W., Miller, D. J., & McDermott, B. (2025). Prevalence of university non-continuation and mental health conditions, and effect of mental health conditions on non-continuation: a systematic review and meta-analysis. *Journal of Mental Health*, 34(2), 222–237. <https://doi.org/10.1080/09638237.2024.2332812>
- Liu, V. Y.-Y., & Lim, S. M. (2022). A psychometric evaluation of the brief resilience scale among tertiary students in Singapore. *Asia Pacific Journal of Education*, 42(3), 464–477. <https://doi.org/10.1080/02188791.2020.1845120>
- Lovibond, S. H. (1995). *Manual for the depression anxiety stress scales*. Sydney Psychology Foundation. <https://cir.nii.ac.jp/crid/1370294643851494273>
- Maizura, N., Pambudi, P. R., Bariyyah, K., Soejanto, L. T., Andrianie, S., & Indreswari, H. (2026). Academic Anxiety Among First-Year Students Living Away From Home During the Transition to Higher Education. *TANJAK: Journal of Education and Teaching*, 7(2), 190–201. <https://doi.org/10.35961/tanjak.v7i2.3018>
- Mauladi, T. A., Hafina, A., Yudha, E. S., & Dynasti, R. H. (2026). Development and Psychometric Properties of a Flourishing Scale for University Students: A Rasch Model Analysis. *Jurnal Pembelajaran, Bimbingan, Dan Pengelolaan Pendidikan*, 6(12), 3. <https://doi.org/10.17977/um065.v6.i12.2026.3>
- Mohamat, R., Sumintono, B., & Hamid, H. (2022). Analisis Kesahan Kandungan Instrumen Kompetensi Guru untuk Melaksanakan Pentaksiran Bilik Darjah Menggunakan Model Rasch Pelbagai Faset (Content Validity Analysis of an Instrument to Measure Teacher's Competency for Classroom Assessment Using Many Facet R. *Jurnal Pendidikan Malaysia*, 47(01), 1–14. <https://doi.org/10.17576/JPEN-2022-47.01-01>
- Mulyanti, S., Sukmawati, W., & Tarkin, N. E. H. (2022). Development of items in Acid-Base Identification Experiments Using Natural Materials: Validity Test with Rasch Model Analysis. *Phenomenon: Jurnal Pendidikan MIPA*, 12(1), 17–30. <https://doi.org/10.21580/phen.2022.12.1.10703>
- Nadhirah, N. A., Ilfiandra Yudha, E. S., Saripah, I Pratiwi, T. I., & Ahman. (2026). Development and psychometric validation of a PERMA-based flourishing scale for university students. *Cypriot Journal of Educational Sciences*, 21(1), 48–62. <https://doi.org/10.18844/cjes.v21i1.9566>
- Nawantara, R. D., Hanurawan, F., Ramli, M., Eva, N., & Amat, S. (2026). Adapting the Psychological Well-Being Scale For Indonesian College Students: A Rasch-Based Validation. *International Journal of Body, Mind and Culture*, 13(2), 139–152. <https://doi.org/10.61838/ijbmc.v13i2.1310>
- Oktasari, M., Hambali, I., Ramli, M., Chusniyah, T., Julian Hikmy, B., Indah Kholili, M., Wahyudin, H., Ariyanto, R. D., Stevani, H., & Hi-yung, L. (2026). Assessing Student Resilience in Indonesia: Validating the Resilience at University Scale (RAUS) Using Rasch Analysis. *Salud, Ciencia Y Tecnología*, 6, 2566. <https://doi.org/10.56294/saludcyt20262566>
- Pointon-Haas, J., Waqar, L., Upsher, R., Foster, J., Byrom, N., & Oates, J. (2024). A systematic review of peer support interventions for student mental health and well-being in higher education. *BJPsych Open*, 10(1), e12. <https://doi.org/10.1192/bjo.2023.603>
- Purnama, D. S., Farozin, M., & Astuti, B. (2022). The Ryff's Psychological Well-Being Scale for Indonesian Higher Education Students: A RASCH Model Analysis. *Indonesian Research Journal in Education | IRJE |*, 6(2), 222–231. <https://doi.org/10.22437/irje.v6i2.20086>
- Quincho, F. S., Rodríguez Galán, D. B., Farfán Pimentel, J. F., Yolanda Josefina, H. F., Arenas, R. D., Crispín, R. L., & Navarro, E. R. (2021). Academic Stress in University Students: Systematic Review. *Ilkogretim Online*, 20(5), 3224. <https://doi.org/10.17051/ilkonline.2021.05.351>
- Ramos-Monsivais, C. L., Rodríguez-Cano, S., Lema-Moreira, E., & Delgado-Benito, V. (2024). Relationship between mental health and students' academic performance through a literature review. *Discover Psychology*, 4(1), 119. <https://doi.org/10.1007/s44202-024-00240-4>
- Ratnaningsih, I. Z., Prihatsanti, U., Prasetyo, A. R., & Sumintono, B. (2025). Validation of the Indonesian

- version of the psychological capital questionnaire (PCQ) in higher education: a Rasch analysis. *Journal of Applied Research in Higher Education*, 17(4), 1243–1258. <https://doi.org/10.1108/JARHE-10-2023-0480>
- Rozental, A., Forsström, D., & Johansson, M. (2023). A psychometric evaluation of the Swedish translation of the Perceived Stress Scale: a Rasch analysis. *BMC Psychiatry*, 23(1), 690. <https://doi.org/10.1186/s12888-023-05162-4>
- Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Sumintono, B., Law, M. Y., & Sitasari, N. W. (2025). Self-handicapping scale: evaluation of psychometric properties among Malaysian and Indonesian university students using Rasch rating scale model analysis. *Journal of Applied Research in Higher Education*, 18(3), 906–920. <https://doi.org/10.1108/JARHE-08-2024-0429>
- Turner, M., & Holdsworth, S. (2024). Supporting a mentally healthy place of study: Examining the relationship between mental well-being, stress, and protective factors among university students. *Scandinavian Journal of Psychology*, 65(5), 816–826. <https://doi.org/10.1111/sjop.13027>
- Wahdi, A. E., Wilopo, S. A., & Erskine, H. E. (2023). The prevalence of adolescent mental disorders in Indonesia: An analysis of Indonesia–National Mental Health Survey (I-NAMHS). *Journal of Adolescent Health*, 72(3), S70. <https://doi.org/10.1016/j.jadohealth.2022.11.143>
- White, A. E., & Karr, J. E. (2025). Psychometric properties of the GAD-7 among college students: Reliability, validity, factor structure, and measurement invariance. *Translational Issues in Psychological Science*, 11(3), 321. <https://doi.org/10.1037/tps0000382>
- Worsley, J., Dryburgh, A., McIntyre, J. C., & Corcoran, R. (2024). Academic and non-academic predictors of common mental health difficulties among university students during the COVID-19 pandemic. *Frontiers in Public Health*, Volume 12. <https://doi.org/10.3389/fpubh.2024.1441176>
- Wulandari, R. M., & Hidayat, R. (2023). Materialism as Personality: Psychometric Properties Using the Rasch Model. *Indonesian Journal of Psychology/Jurnal Psikologi*, 50(2), 157. <https://doi.org/10.22146/jpsi.77592>
- Yusuf, S., Budiman, N., Saripah, I., Afendy, P. M., Gofur, R., & Hamidah, N. H. (2025). Development and rasch validation of the student mental health scale among Indonesian university students. *ProGCouns: Journal of Professionals in Guidance and Counseling*, 6(1), 40–50. <https://doi.org/10.21831/progcouns.v6i1.79287>
- Zhai, Y., & Carney, J. V. (2024). The role of mental health and protective factors in student academic persistence and retention during a global crisis. *Cambridge Prisms: Global Mental Health*, 11, e15. <https://doi.org/10.1017/gmh.2024.12>