Psychometric Validation of the Mental Toughness Questionnaire-18 Among High School Students Using the Rasch Model Analysis

Akbar Gibran Yusri, Arbin Janu Setiyowati*, Diniy Hidayatur Rahman

Universitas Negeri Malang, Malang, Indonesia
*Corresponding author, email: arbin.janu.fip@um.ac.id

https://doi.org/10.17977/um065.v7.i5.2027.2

Article history
Submitted: 24 July 2026
Revised: 3 October 2026
Accepted: 23 October 2026
Published: 29 October 2026

Keywords
High school students
Mental toughness
MTQ-18
Psychometric validation
Rasch model

Abstract
The increasing academic, social, and developmental demands faced by high school students make mental toughness an important psychological resource for adapting to various pressures. Therefore, measuring mental toughness is essential to support needs-based guidance and counseling services. This study aimed to examine the psychometric characteristics of the Mental Toughness Questionnaire-18 (MTQ-18) among general high school students using the Rasch Model. A quantitative approach with a cross-sectional survey design was employed. The participants consisted of 455 students in SIDRAP Regency who were selected using convenience sampling. The instrument used was the Indonesian version of the MTQ-18, which had been adapted to the characteristics of high school students. Data were analyzed using the Rasch Model with Winsteps software. The findings indicate that the MTQ-18 provides preliminary evidence of adequate psychometric quality for measuring mental toughness among high school students. The five-point Likert response categories functioned consistently in differentiating respondents' choices. Most items demonstrated acceptable fit with the measurement model, although several item indicators require further evaluation. The instrument also met the unidimensionality assumption, with a raw variance explained by measures value of 47.2%. Item difficulty analysis showed variations in item characteristics across different levels of students' mental toughness. Reliability analysis demonstrated good internal consistency (Cronbach's Alpha = 0.72), very high item reliability (1.00), and adequate person reliability (0.68). These findings suggest that the MTQ-18 has potential as an initial assessment instrument (need assessment) to support guidance and counseling services for high school students.

1. Introduction

The psychological well-being of students is a key concern in 21st-century education, particularly amid increasing academic demands, social pressures, competition, and the complexities of adolescent development (Boonma, 2025). At the high school level, students are in a developmental phase that requires them not only to achieve academic success but also to adapt socially and begin making decisions regarding their education and careers (Flores-Builes & Mateu-Pérez, 2025). At the same time, students also face the possibility of failure, academic stress, and various pressures that arise during the educational process (Deng et al., 2022). These conditions require students to have the ability to adapt and continue to function optimally when facing various pressures (Singh & Bisht, 2025). One psychological construct that can explain this ability is mental toughness.

Mental toughness was initially developed in the field of sports psychology to describe the psychological characteristics that enable individuals to maintain their performance in competitive and high-pressure situations. Clough, Earle, and Sewell, (2002) define mental toughness as an individual's ability to remain confident, focused, self-controlled, and committed when facing challenges and obstacles. This concept later evolved into a set of psychological characteristics that influence an individual's ability to respond adaptively to pressure, challenges, and stress without losing sight of their goals (Strycharczyk, Clough, & Perry, 2021). Guszkowska and Wójcik, (2021) also emphasize that mental toughness is related to an individual's ability to persevere and remain goal-oriented even in stressful situations. Therefore, mental toughness is viewed as one of the psychological resources that help individuals cope with difficult and challenging situations (Hsieh et al., 2024), including life circumstances that require the ability to adapt to various pressures (Kiarostami, Jeddi, Roohinezhad, & Hakimi, 2022).

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Conceptually, mental toughness is developed through the 4C model, which encompasses control, commitment, challenge, and confidence (Strycharczyk et al., 2021). Control describes an individual's belief in their ability to manage their emotions, life, and responses to the situations they face (Bandura, 2023). Commitment relates to perseverance in sustaining effort, completing tasks, and achieving set goals (Lee & Kim, 2023). Challenge refers to an individual's tendency to view obstacles, changes, and pressures as opportunities for growth (Lee & Kim, 2023). Meanwhile, confidence describes an individual's belief in their ability to complete tasks and face challenges effectively (Jiang, Fu, & Dong, 2025). These four aspects demonstrate that mental toughness is not only related to the ability to persevere in difficult situations but also encompasses the ability to exercise self-control, maintain effort, respond positively to challenges, and believe in one's own capabilities.

Although it initially developed in the context of sports, mental toughness is now increasingly studied across various fields, including education, organizations, mental health, and adolescent development (Farnsworth et al., 2021). In the context of education, mental toughness is relevant because students also face various demands that require the ability to manage emotions, maintain motivation, build psychological resilience, and adapt effectively (Janjani et al., 2025). This construct is known to be associated with academic resilience and coping abilities in the face of academic pressure (Clair-thompson & Devine, 2023). Furthermore, mental toughness is also linked to students' psychological well-being, learning motivation, and academic achievement (S. Zhang, Li, & Unger, 2022). Students with higher levels of mental toughness tend to be better able to cope with academic pressure, manage failure, and maintain optimism throughout their learning (Fengjun, Jing, & Liguang, 2022). The growing body of research on mental toughness in an educational context indicates that this construct is not only relevant to athletes but is also crucial for understanding the psychological characteristics of students (Yusri, Indreswari, Bariyyah, Setyowati, & Rahman, 2026). Therefore, a measurement tool capable of accurately assessing students' mental toughness is needed.

One of the most widely used instruments for measuring mental toughness is the Mental Toughness Questionnaire (MTQ), which was developed based on the 4C model. One of its early versions was the MTQ-48, consisting of 48 items, which has been used in various studies on mental toughness (Perry et al., 2021). However, the relatively large number of items can be a consideration when using the instrument in field research, especially when administered to school-aged respondents. To produce a more concise and practical measurement, the Mental Toughness Questionnaire-18 (MTQ-18), consisting of 18 items, was developed (Dagnall et al., 2019). This shorter version offers efficiency in the measurement process; however, the application of an instrument to different populations and contexts still requires psychometric testing to ensure that its measurement quality is adequately established.

A number of studies have examined the use and psychometric characteristics of mental toughness instruments across various populations. (Dagnall et al., 2019) evaluated the MTQ-18 among students aged approximately 16 years and provided evidence of the instrument's applicability to school-aged populations. However, that study was conducted in an educational context outside Indonesia and thus does not provide insight into the measurement characteristics of the MTQ-18 among Indonesian students. Haqiyah et al., (2023) also used the Rasch Model approach in the development of a mental toughness instrument, but their study focused on a population of athletes rather than the MTQ-18. Furthermore, Denovan, Dagnall, Artamonova, and Papageorgiou, (2024) evaluated a mental toughness instrument using the Rasch Model; however, the instrument analyzed was the MTQ-10, so the results do not directly reflect the psychometric characteristics of the MTQ-18. In the Indonesian context, Yudiarso, Ardhiani, Surya, Watimena, and Kanzaki, (2025) tested the Indonesian version of the MTQ-18 using the Rasch Model, but the participants were drawn from a population of athletes. Thus, previous studies have provided important evidence regarding the use of the MTQ-18 with students as well as Rasch-based psychometric evaluations; however, each study still differs in terms of cultural context, sample type, instrument, and population characteristics compared to the use of the MTQ-18 with general high school students in Indonesia.

Based on previous research, psychometric evidence for the Rasch Model-based MTQ-18 among general high school students or non-athletes in the Indonesian educational context remains limited. Evidence regarding school-aged students was previously obtained from educational contexts outside Indonesia. Meanwhile, testing of the Indonesian version of the MTQ-18 using the Rasch Model has so far been conducted only on athletic populations. This situation highlights the need to understand how the MTQ-18 functions when applied to general high school students in Indonesia, who have distinct developmental contexts and academic demands. The Rasch model is relevant for addressing this need because it allows for the evaluation of instrument quality at both the item and respondent levels (Engelhard Jr & Wang, 2024), including response category function, item fit, item difficulty, reliability, and instrument unidimensionality (Rumtini et al., 2026). This approach is also appropriate when an instrument is administered to a population different from the context of previous testing because the characteristics of each item can be evaluated more specifically (Hadžibajramović, Schaufeli, & De Witte, 2022).

Based on the above description, this study aims to examine the psychometric characteristics of the Mental Toughness Questionnaire-18 (MTQ-18) among general high school students in Indonesia using the Rasch Model.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

The analysis focuses on response category function, item fit, unidimensionality, item difficulty, and instrument reliability. This study is expected to expand the psychometric evidence for the MTQ-18 in the context of Indonesian education, particularly among non-athlete high school students. The results are also expected to serve as a foundation for further testing and to support the judicious use of the MTQ-18 as a needs assessment instrument in guidance and counseling services.

2. Method

2.1. Research Design

The MTQ-18 scale has previously been widely used among athletes; therefore, this study adapted the context and language to align with the characteristics of high school students. This study used the Rasch Model to analyze the psychometric quality of the instrument at the item and respondent levels (Rumtini et al., 2026). The use of this approach aimed to produce a valid and reliable mental toughness scale for use among high school students in general.

2.2. Participants

The study population consisted of high school students in Sidenreng Rappang (Sidrap) Regency, South Sulawesi, Indonesia. Participants were selected using convenience sampling, a non-probability sampling technique, which involves selecting participants based on ease of access, availability, and willingness to participate in the study (Golzar & Noor, 2022). The study was conducted at two high schools, whose identities were disguised as School A and School B to maintain the confidentiality of the institutions. Both schools were selected based on the researcher's ease of access and the schools' willingness to grant permission to conduct the study. This selection allowed the researcher to coordinate directly with school officials regarding the distribution of survey instruments and the collection of student data. The questionnaire was administered via Google Forms and completed by students through direct coordination with school officials. The data collection process yielded 455 participants, the total number of complete responses successfully obtained from both schools during the data collection period and this number met the data completeness requirements for analysis. The distribution of participants based on demographic characteristics is presented in Table 1.

Table 1. Distribution of the Research Sample Based on Demographic Characteristics

Demographic Characteristics	Category	School A	School B	Total
Gender	Male	65	61	126
	Female	142	187	329
Total		207	248	455
Age	15 Years old	54	21	75
	16 Years old	83	77	160
	17 Years old	67	91	158
	18 Years old	3	59	62
Total		207	248	455
Grade Level	X	106	90	196
	XI	95	48	143
	XII	6	110	116
Total		207	248	455

2.3. Instrument

The instrument used in this study was the Indonesian-language version of the Mental Toughness Scale developed by Yudiarso et al., (2025). This instrument is an adaptation of the Mental Toughness Questionnaire (MTQ-18) developed by Dagnall et al., (2019), which is a shortened version of the Mental Toughness Questionnaire (MTQ-48) developed by Peter Clough. During the adaptation process, Yudiarso et al., (2025) carried out a series of steps, including translation, expert judgment, cultural and linguistic adaptation, and a pilot test to ensure equivalence of meaning and appropriateness of context for use among the Indonesian population.

The MTQ-18 scale consists of 18 favorable and unfavorable statements that measure the four main dimensions of mental toughness: control, commitment, challenge, and confidence. Each item is structured using a five-point Likert scale: 1 (Strongly Disagree), 2 (Disagree), 3 (Neutral), 4 (Agree), and 5 (Strongly Agree). In a previous study, this scale was pilot-tested on 400 participants athletes in Indonesia aged 13–56 years. However, in this study, the use of the instrument was focused on high school students in general; therefore, several adjustments were made to the wording to make it easier for high school students to understand, while still adhering to the original structure of the MTQ-18, which consists of 18 items and four main dimensions.

In the process of adapting the wording of the items, the researchers utilized generative artificial intelligence to explore alternative word choices and simpler sentence structures appropriate for the comprehension level of high school students. The use of AI was limited to editorial aspects and was not used to

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

alter the constructs, indicators, or psychological substance of the items. Each alternative wording was reviewed by the researchers with reference to the original meaning of the item. Next, content equivalence between the items before and after the adaptation was reviewed by guidance and counseling experts to ensure that the wording changes did not alter the meaning or the mental toughness construct being measured. These adaptations were made to improve readability without changing the indicators or the psychological substance of the mental toughness construct. The instrument blueprint is presented in Table 2.

Table 2. Blueprint of the Mental Toughness Scale with Adjustments for High School Students in General

Dimensions	MTQ-18 Indonesian Version	MTQ-18 Adapted for High School Students	MTQ-18 Adapted for High School Students in English version	Statement Items	Item
Control	*Bahkan ketika berada dibawah tekanan yang besar, saya biasanya tetap tenang*	*Bahkan ketika berada dibawah tekanan yang besar, saya biasanya tetap tenang*	Even when under a lot of pressure, I usually stay calm	(+)	Q1
	Saya cenderung mengkhawatirkan segala sesuatunya jauh sebelum hal itu terjadi	*Saya cenderung mengkhawatirkan segala sesuatu jauh sebelum hal itu terjadi*	I tend to worry about everything long before it happens	(-)	Q2
	Saya biasanya merasa memegang kendali	*Saya biasanya merasa memegang kendali*	I usually feel in control	(+)	Q10
	Ketika lelah, saya merasa kesulitan untuk memulai sesuatu	*Saat saya lelah, saya merasa kesulitan untuk memulai sesuatu*	When I'm tired, I find it hard to get started on something	(-)	Q12
	Saya biasanya sulit untuk rileks	*Saya biasanya sulit untuk rileks*	I usually have a hard time relaxing	(-)	Q16
Commitment	*Saya biasanya sulit membangkitkan semangat terhadap tugas-tugas yang harus dikerjakan*	*Saya biasanya sulit membangkitkan semangat terhadap tugas-tugas yang harus dikerjakan*	I usually have a hard time getting motivated to tackle the tasks at hand	(-)	Q3
	"Saya tidak tahu harus mulai dari mana" adalah perasaan yang biasa saya rasakan ketika dihadapkan pada beberapa hal yang harus dilakukan sekaligus	*"Saya tidak tahu harus mulai dari mana" adalah perasaan yang biasa saya rasakan ketika dihadapkan pada beberapa hal yang harus dilakukan sekaligus*	"I don't know where to start" is a feeling I often have when faced with several things to do at once	(-)	Q6
	Saya biasanya merasa sulit untuk melakukan usaha mental ketika Lelah	*Saya biasanya merasa sulit untuk berpikir atau fokus ketika lelah*	I usually find it hard to think or focus when I'm tired	(-)	Q17
Challenge	*Saya biasanya mengatasi dengan baik setiap masalah yang terjadi*	*Saya biasanya mampu mengatasi dengan baik setiap masalah yang terjadi*	I'm usually able to handle any problem that arises well	(+)	Q4
	Saya berharap hidup saya lebih bisa diprediksi	*Saya berharap hidup saya lebih mudah diprediksi*	I wish my life were more predictable	(-)	Q11
	Saya umumnya mampu bereaksi dengan cepat saat sesuatu yang tidak terduga terjadi	*Saya umumnya mampu bereaksi dengan cepat saat sesuatu yang tidak terduga terjadi*	I'm generally able to react quickly when something unexpected happens	(+)	Q13
Confidence	*Secara umum, saya merasa bahwa saya adalah orang yang berharga*	*Secara umum, saya merasa bahwa saya adalah orang yang berharga*	In general, I feel that I am a valuable person	(+)	Q5
	Saya selalu berbicara jujur ketika memiliki sesuatu untuk disampaikan	*Saya selalu mengungkapkan pendapat ketika memiliki sesuatu untuk disampaikan*	I always speak up when I have something to say	(+)	Q7
	Ketika saya membuat kesalahan, saya cenderung merasa khawatir selama beberapa hari	*Ketika saya membuat kesalahan, saya cenderung merasa khawatir selama beberapa hari*	When I make a mistake, I tend to worry about it for several days	(-)	Q8
	Dalam diskusi, saya sering mundur meskipun merasa yakin tentang sesuatu	*Dalam diskusi, saya sering mengalah meskipun*	In discussions, I often back down even when I'm	(-)	Q9

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Dimensions	MTQ-18 Indonesian Version	MTQ-18 Adapted for High School Students	MTQ-18 Adapted for High School Students in English version	Statement Items	Item
		merasa yakin tentang sesuatu	certain about something		
	Meskipun keadaannya buruk, saya yakin semuanya akan berakhir baik	*Seburuk apapun keadaannya, saya yakin semuanya akan berakhir baik*	No matter how bad things get, I'm confident everything will turn out okay	(+)	Q14
	Saya biasanya melihat sisi positif kehidupan	*Saya biasanya melihat sisi positif kehidupan*	I usually look on the bright side of life	(+)	Q15
	Jika saya merasa seseorang salah, saya tidak takut untuk berdebat dengan mereka	*Jika saya merasa seseorang salah, saya tidak takut untuk berdebat dengan mereka*	If I think someone is wrong, I'm not afraid to argue with them	(+)	Q18

2.4. Data Analysis

Data analysis in this study was conducted using the Rasch Model approach to test the psychometric quality of the mental toughness scale, in terms of construct structure, item characteristics, and respondent characteristics. The analysis was performed using Winsteps software, version 3.73, to obtain a more comprehensive evaluation of the instrument. The Rasch model was used to analyze item fit, respondent ability, instrument reliability, and the functioning of categories on the rating scale (Tesio, Caronni, Kumbhare, & Scarano, 2024).

The first stage of the Rasch analysis is the rating scale analysis, which aims to evaluate respondents' ability to accurately distinguish between each response category on a 1-to-5 Likert scale. The second stage involves analyzing item fit and validity to determine whether each item functions in accordance with the measurement model, based on the indicators Mean Square of Fit (MNSQ), Z-Standard of Fit (ZSTD), and Point Measure Correlation (Pt Measure Corr). Next, the third stage tests the assumption of unidimensionality of the instrument through the raw variance explained by measures and the unexplained variance in the first through fifth contrasts of residuals. The fourth stage involved analyzing item difficulty levels based on logit values to determine the variation in difficulty levels for each item. Additionally, the analysis continued with the presentation of the Wright Person-Item Map to examine the overlap between the distribution of respondent ability (person ability) and item difficulty levels on the same logit scale. The fifth stage involved a Differential Item Functioning (DIF) analysis to determine whether each item functioned equivalently across different respondent groups based on gender and to identify potential differences in item difficulty between male and female students. The final stage involved an analysis of summary statistics to obtain information regarding the quality of the instrument, including internal consistency (Cronbach's Alpha), person reliability, item reliability, and item-response separation indices.

3. Results and Discussion

The results of the research data analysis include an evaluation of the instrument's quality based on the Rasch Model, which encompasses an analysis of the rating scale, item fit, unidimensionality tests, item difficulty levels, DIF analysis, and instrument reliability.

3.1. Rating Scale Diagnostic

A rating scale analysis was conducted to determine the extent to which participants understood and distinguished between the response options available in the instrument (Chong, Mokshein, & Mustapha, 2022). In the Rasch Model, this analysis is used to evaluate whether the response categories on the Likert scale function properly and are used consistently by respondents. Respondents' understanding of the five response categories can be considered good if the observed average and Andrich threshold values show a pattern of gradual increase from the lowest to the highest category (Ismail et al., 2024). The observed average and Andrich threshold values are shown in Table 3.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Table 3. Results of the Rating Scale Diagnostic Analysis

Category Label	Score	Observed Count	%	Observed Average	Sample Expect	Infit MNSQ	Outfit MNSQ	Andrich Threshold	Category Measure	
1	1	931	11	-1.57	-1.57	1.02	1.02	NONE	(-3.26)	1
2	2	1865	23	-.96	-.89	.87	.88	-1.93	-1.55	2
3	3	3144	38	-.01	-.07	.95	.93	-1.01	.00	3
4	4	1621	20	.71	.70	1.00	.99	1.00	1.55	4
5	5	629	8	1.19	1.29	1.12	1.13	1.95	(3.27)	5

Based on the results of the rating scale analysis, it was found that the observed average score increased gradually from category 1 to category 5, namely -1.57; -0.96; -0.01; 0.71; and 1.19. This increase indicates that respondents were able to consistently distinguish between response options from the lowest to the highest level. In addition, the Andrich Threshold values also showed a sequentially increasing pattern, namely -1.93 in category 2, -1.01 in category 3, 1.00 in category 4, and 1.95 in category 5. The first category had no threshold value because it served as the starting category. This pattern of increase indicates that there are no disruptions or irregularities in the structure of the response categories. Thus, the five-category Likert scale used in this mental toughness scale can be said to function well and be understood by the respondents.

3.2. Item Fit and Validation

Item fit is used to assess whether each item of the instrument functions optimally in measuring mental toughness so as not to cause misunderstandings among respondents (Medvedev & Krägeloh, 2022). This evaluation was conducted based on the results of data analysis using Winsteps. According to the study by Amiruddin, (2026) item fit assessment is performed using three main indicators: Outfit Mean Square (MNSQ), Outfit Z-Standard (ZSTD), and Point Measure Correlation (Pt Measure Corr).

An item is deemed acceptable if its Mean Square Error (MNSQ) falls within the range of 0.5 to 1.5 ($0.5 < MNSQ < 1.5$). Furthermore, an item is also acceptable if its Z-Standard (ZSTD) falls within the range of -2.0 to +2.0 ($-2.0 < ZSTD < +2.0$). As for the Point Measure Correlation indicator, an item is accepted if it has a value between 0.40 and 0.85 ($0.40 < Pt\ Measure\ Corr < 0.85$). A statement item can be categorized as a "fit" if it meets at least one of the three specified fit indicators, meaning the item is still considered suitable for use in the measurement instrument (Norhayati, Imtithal, & Akma, 2021). The results of the item misfit analysis are presented in Table 4.

All items had Outfit Mean Square (MNSQ) values ranging from 0.5 to 1.5. These results indicate that no items were outliers based on the Outfit MNSQ criteria; thus, in general, all items remained within the acceptance limits of the Rasch model.

However, based on the Outfit Z-Standard (ZSTD) indicator, several items were found to have values outside the criterion range of -2.0 to +2.0, namely items P11, P5, P18, P9, P2, P10, P17, P13, and P4. Excessively high ZSTD values indicate that respondents' response patterns to these items do not align well with the measurement model's expectations. Conversely, excessively low ZSTD values suggest that response patterns to these items tend to be too consistent or too predictable compared to what the Rasch model expects.

Furthermore, regarding the Point Measure Correlation (Pt Measure Corr) indicator, several items had values below 0.40, namely P11, P18, P8, P14, P1, P15, P10, and P13. These values indicate that the relationship between the items and the construct being measured is relatively weak. However, since each item still meets at least one of the three established item fit criteria, all items are categorized as "fit" and deemed suitable for measuring mental toughness. Thus, all items in the instrument are deemed capable of functioning adequately without causing respondents to misinterpret the statements provided.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Table 4. Results of the Item Misfit Analysis

Entry Number	Total Score	Total Count	Measure	Model S.E.	Infit		Outfit		Pt-Measure		Exact	Exact	Item
					MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Exp.	Obs%	Exp%	
11	1102	455	.74	.06	1.37	5.4	1.37	5.4	A.18	.41	42.4	44.6	P11
5	1525	455	-.71	.06	1.23	3.3	1.23	3.4	B.46	.42	42.9	47.0	P5
18	1619	455	-1.07	.06	1.20	3.1	1.20	3.0	C.33	.42	38.0	44.1	P18
9	1158	455	.55	.06	1.17	2.6	1.17	2.6	D.48	.42	43.5	46.3	P9
8	827	455	1.73	.06	1.10	1.7	1.07	1.2	E.38	.38	47.0	45.1	P8
14	1847	455	-1.85	.06	1.09	1.5	1.07	1.2	F.39	.40	44.0	43.6	P14
6	1011	455	1.05	.06	1.05	.9	1.05	.9	G.47	.41	43.1	43.1	P6
1	1507	455	-.65	.06	1.00	.0	1.00	.1	H.32	.42	49.9	47.8	P1
3	1280	455	.13	.06	.93	-1.0	.95	-.8	I.45	.42	48.1	49.6	P3
15	1692	455	-1.29	.06	.92	-1.3	.94	-1.1	i.35	.41	46.2	43.8	P15
16	1247	455	.24	.06	.93	-1.1	.92	-1.2	h.53	.42	53.8	49.1	P16
12	1273	455	.10	.06	.91	-1.5	.92	-1.4	g.51	.41	48.8	43.3	P12
7	1544	455	-.78	.06	.91	-1.4	.91	-1.4	f.47	.42	53.0	46.8	P7
2	993	455	1.44	.06	.90	-1.9	.88	-2.1	e.45	.39	51.2	43.7	P2
10	1300	455	.06	.06	.86	-2.1	.86	-2.1	d.37	.42	51.9	50.0	P10
17	959	455	1.24	.06	.85	-2.6	.84	-2.8	c.55	.40	45.7	43.3	P17
13	1582	455	-.91	.06	.83	-2.8	.83	-2.8	b.32	.42	49.0	45.1	P13
4	1586	455	-.92	.06	.64	-6.5	.64	-6.5	a.44	.42	57.8	45.1	P4
Mean	1317.9	455.0	.00	.06	.99	-.2	.99	-.3			47.6	45.6	
SD	297.2	.0	1.03	.00	.17	2.7	.17	2.7			4.8	2.2	

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

3.3. Unidimensionality

The unidimensionality of an instrument is one of the key indicators for assessing whether the developed instrument is capable of measuring the construct it is intended to measure (Chiu et al., 2024). In Rasch Model analysis, unidimensionality is tested using Principal Component Analysis (PCA) on the residuals to determine the extent to which the instrument is able to represent the single main construct being measured (Cioffi, Galano, Passeggia, & Testa, 2024; Tennant & Küçükdeveci, 2023). The instrument is considered unidimensional if the raw variance explained by measures reaches ≥ 20%. Generally, the empirical unidimensionality value is categorized as adequate if it falls within the 20–40% range, good if within the 40–60% range, and very good if it exceeds 60% (Latifah, Saripah, Dodi, & Sunarya, 2024). Additionally, a measurement can be considered unidimensional if the Unexplained Variance in the 1st to 5th Contrast of Residuals is below 15% for each contrast (Lee & Lee, 2025; Muslihin, Suryana, Ahman, Suherman, & Dahlan, 2022). The results of the unidimensionality analysis are presented in Table 5.

Table 5. Result of the Item Unidimensionality Analysis

	Empirical			Modeled
Total raw variance in observations	34.1	100.0%	-	100.0%
Raw variance explained by measures	16.1	47.2%	-	47.0%
Raw variance explained by persons	2.3	6.8%	-	6.8%
Raw variance explained by items	13.8	40.4%	-	40.2%
Raw unexplained variance (total)	18.0	52.8%	100.0%	53.0%
Unexplained variance in 1st contrast	3.8	11.1%	21.0%	-
Unexplained variance in 2nd contrast	1.5	4.3%	8.1%	-
Unexplained variance in 3rd contrast	1.4	4.2%	7.9%	-
Unexplained variance in 4th contrast	1.2	3.5%	6.6%	-
Unexplained variance in 5th contrast	1.1	3.2%	6.1%	-

The results of the unidimensionality analysis show that the raw variance explained by measures is 47.2%. This value meets the minimum criterion of ≥ 20% and falls into the "good" category because it lies within the 40–60% range. Additionally, the unexplained variance values for the 1st through 5th contrasts were 11.1%, 4.3%, 4.2%, 3.5%, and 3.2%, respectively. All of these values are below the 15% threshold, so there is no strong indication of secondary dimensions interfering with the main construct. Thus, the mental toughness scale can be said to meet the assumption of unidimensionality and is suitable for measuring the intended construct.

3.4. Item Difficulty Level

Item difficulty is an indicator that shows how easy or difficult an item on a test is for respondents (Abdellatif, 2023). In the Rasch Model, item difficulty is determined based on the logit measure, taking into account the standard deviation (SD) of the item's logit value. The higher an item's logit value, the higher its difficulty level; conversely, a lower logit value indicates that the item is relatively easier for respondents. According to Muqaffi, Rusmana, and Ilfiandra, (2024), item difficulty levels are classified into four categories: very difficult if the logit measure is greater than +1 SD; difficult if the logit measure is between 0 and +1 SD; moderate if the logit measure is between −1 SD and 0; and easy if the logit measure is less than −1 SD. This categorization was used as the basis for analyzing the distribution of difficulty levels for each item in the instrument. The results of the item difficulty analysis based on these four categories are presented in Table 6.

Based on the results of the item difficulty analysis, the standard deviation of the measure was found to be 1.03 logits, with an average measure of 0.00 logits. Thus, based on the categorization formula explained earlier, items with a measure value greater than 1.03 are categorized as very difficult; items with a measure value between 0.00 and 1.03 are categorized as difficult; items with a measure value between -1.03 and less than 0.00 are categorized as moderate; and items with a measure value less than -1.03 are categorized as easy. The distribution of item difficulty levels is presented in Table 7.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Table 6. Results of the Item Difficulty Analysis

Entry Number	Total Score	Total Count	Measure	Model S.E.	Infit		Outfit		Pt-Measure		Exact	Exact	Item
					MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Exp.	Obs%	Exp%	
8	827	455	1.73	.06	1.10	1.7	1.07	1.2	.38	.38	47.0	45.1	P8
2	903	455	1.44	.06	.90	-1.9	.88	-2.1	.45	.39	51.2	43.7	P2
17	959	455	1.24	.06	.85	-2.6	.84	-2.8	.55	.40	45.7	43.3	P17
6	1011	455	1.05	.06	1.05	.9	1.05	.9	.47	.41	43.1	43.1	P6
12	1023	455	1.01	.06	.91	-1.5	.92	-1.4	.51	.41	48.8	43.3	P12
11	1102	455	.74	.06	1.37	5.4	1.37	5.4	.18	.41	42.4	44.6	P11
9	1158	455	.55	.06	1.17	2.6	1.17	2.6	.48	.42	43.5	46.3	P9
16	1247	455	.24	.06	.93	-1.1	.92	-1.2	.53	.42	53.8	49.1	P16
3	1280	455	.13	.06	.93	-1.0	.95	-.8	.45	.42	48.1	49.6	P3
10	1300	455	.06	.06	.86	-2.1	.86	-2.1	.37	.42	51.9	50.0	P10
1	1507	455	-.65	.06	1.00	.0	1.0	.1	.32	.42	49.9	47.8	P1
5	1525	455	-.71	.06	1.23	3.3	1.23	3.4	.46	.42	42.9	47.0	P5
7	1544	455	-.78	.06	.91	-1.4	.91	-1.4	.37	.42	53.0	46.8	P7
13	1584	455	-.91	.06	.83	-2.8	.83	-2.8	.42	.42	49.0	45.1	P13
4	1586	455	-.92	.06	.64	-6.5	.64	-6.5	.34	.42	57.8	45.1	P4
18	1629	455	-1.07	.06	1.20	3.1	1.20	3.0	.33	.42	38.0	44.1	P18
15	1692	455	-1.29	.06	.92	-1.3	.94	-1.1	.35	.41	46.2	43.8	P15
14	1847	455	-1.85	.06	1.09	1.5	1.07	1.2	.39	.40	44.0	43.6	P14
Mean	1317.9	455.0	.00	.06	.99	-.2	.99	-.3			47.6	45.6	
SD	297.2	.0	1.03	.06	.17	2.7	.17	2.7			4.8	2.6	

Table 7. Distribution of Item Difficulty Levels

Criteria	Number of Items	Items
Very difficult	4 items	P8, P2, P17, P6
Difficult	6 items	P12, P11, P9, PP16, P3, P10
Moderate	5 items	P1, P5, P7, P13, P4
Easy	3 items	P18, P15, P14

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

The categorization results show that there are 4 items in the "very difficult" category, 6 items in the "difficult" category, 5 items in the "moderate" category, and 3 items in the "easy" category. This distribution indicates that the MTQ-18 has a fairly wide range of item difficulty levels. This diversity in difficulty levels suggests that the instrument is not only capable of measuring students with a specific level of mental toughness but can also provide information on students with both lower and higher abilities.

To determine the extent to which item difficulty levels can cover the range of students' mental toughness abilities, a further analysis was conducted using the Wright Person-Item Map (Folastri, Hambali, Ramli, Sa'dun Akbar, & Sofyan, 2023). This analysis maps the distribution of respondents' abilities (person measure) and item difficulty levels (item measure) onto the same logit scale, thereby revealing the alignment between respondents' abilities and the characteristics of the available items.

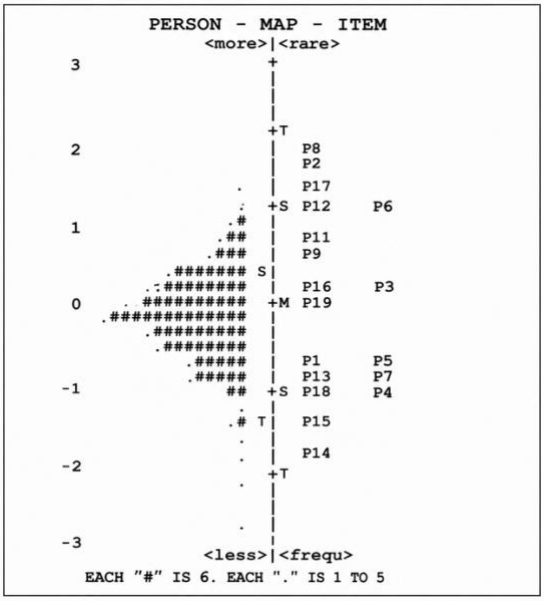


Figure 1. Person-Item Map of the Mental Toughness Questionnaire-18 (MTQ-18)

Based on the Wright Person-Item Map in the figure, the distribution of students' abilities spans a relatively wide range, with most respondents clustered around the construct mean (0 logit). Meanwhile, the items are distributed across a range of difficulty levels, from low to high, with item P14 being the easiest and P8 the most difficult. Most of the distribution of students' abilities falls within the range of item difficulty levels available. This indicates that the MTQ-18 has sufficient measurement coverage to identify variations in the levels of mental toughness among high school students. The available items are capable of capturing most respondent characteristics, including students with low, moderate, and high levels of mental toughness. However, it appears that a small proportion of respondents have levels of mental toughness that are higher than the highest-difficulty item or lower than the easiest item. This suggests that the instrument could still be improved by adding items with more extreme difficulty levels to optimize measurement for respondents with extreme characteristics.

Overall, the results of the item difficulty analysis and the Wright Person-Item Map indicate that the MTQ-18 has a sufficiently good item distribution for measuring the construct of mental toughness among high school students. The instrument features a range of difficulty levels capable of covering most of the respondents' ability spectrum, thereby supporting the use of the MTQ-18 as a measurement instrument for mental toughness in an educational context.

3.5. Differential Item Functioning (DIF)

A Differential Item Functioning (DIF) analysis was conducted to determine whether each item on the instrument functions as an equivalent measure across different groups of respondents (Tesio, Caronni, Simone, Kumbhare, & Scarano, 2024). In this study, the DIF analysis was conducted based on gender to determine whether the MTQ-18 items had different levels of difficulty between male and female students even though they had relatively similar levels of mental toughness. DIF identification was based on the probability (PROB) values in the DIF analysis results table. Items with a PROB value < 0.05 were categorized as exhibiting DIF because they indicated a significant difference in item function between groups, whereas items with a PROB value > 0.05 indicated that the items functioned relatively equivalently across the compared groups (Tennant & Küçükdeveci, 2023). A summary of the DIF analysis results is presented in Table 8.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Table 8. Summary of Differential Item Functioning (DIF) Analysis Results

Item	Summary DIF CHI-SQUARE	D.F	PROB	BETWEEN-CLASS MEAN-SQUARE	t=ZSTD	Indication for DIF
1	5.0615	1	0.0245	2.0388	1.0400	DIF
2	0.9040	1	0.3417	0.3968	-0.0911	No
3	2.8420	1	0.0918	1.1385	0.5651	No
4	0.7306	1	0.3927	0.3052	-0.2217	No
5	8.5976	1	0.0034	3.5519	1.5867	DIF
6	0.4415	1	0.5064	0.1933	-0.4234	No
7	0.1267	1	0.7219	0.0608	-0.8157	No
8	2.9546	1	0.0856	1.2864	0.6572	No
9	2.1129	1	0.1461	0.8373	0.3495	No
10	0.5355	1	0.4643	0.2176	-0.3740	No
11	0.1715	1	0.6788	0.0811	-0.7316	No
12	1.5937	1	0.2068	0.6753	0.2112	No
13	0.7072	1	0.4004	0.2955	-0.2369	No
14	5.7974	1	0.0161	2.3515	1.1710	DIF
15	0.0000	1	1.0000	0.0000	-1.2642	No
16	0.0000	1	1.0000	0.0007	-1.4638	No
17	6.8039	1	0.0091	2.8941	1.3731	DIF
18	0.0397	1	0.8422	0.0153	-1.1233	No

Based on the DIF analysis results in Table 9, most MTQ-18 items demonstrate relatively equivalent measurement functioning between male and female students. Of the 18 items analyzed, 14 had a probability (PROB) value > 0.05, indicating no evidence of DIF. This suggests that most items consistently measure the construct of mental toughness across both gender groups. However, there were four items with probability values < 0.05, namely items P1 (PROB = 0.0245), P5 (PROB = 0.0034), P14 (PROB = 0.0161), and P17 (PROB = 0.0091). Statistically, these four items indicate a difference in function based on gender. Nevertheless, the DIF findings do not directly imply that these items should be eliminated; rather, they indicate that the items require further review, taking into account the magnitude of the difference (DIF contrast) and the item's substantive fit to the construct of mental toughness.

To determine the direction and magnitude of item functioning differences between male and female students, a pairwise DIF analysis was conducted (Albano, French, & Vo, 2024). This analysis provides information regarding differences in item difficulty (DIF Measure) as well as DIF Contrast values between the two groups. The results of the pairwise DIF analysis are presented in Table 9.

Based on the pairwise DIF analysis results, the four items previously identified as exhibiting DIF showed different directions of functional differences between male and female students. Item P1 had a DIF contrast value of -0.30 logits, indicating that the item was relatively more difficult for female students than for male students. Conversely, item P5 has a DIF contrast value of 0.39 logits, indicating that the item is relatively more difficult for male students. Furthermore, item P14 has a DIF contrast value of 0.33 logits, indicating a tendency for the item to be more difficult for male students, while item P17 has a DIF contrast value of -0.35 logits, indicating a tendency for the item to be more difficult for female students. Although these four items showed statistically significant differences in function, the DIF contrast values fell within a relatively narrow range, between 0.30 and 0.39 logits. This indicates that the magnitude of the differences in item function between gender groups remains relatively low, suggesting no significant deviation in item function.

The DIF plot in Figure 2 shows the pattern of MTQ-18 item difficulty levels between male and female student groups. Overall, most item positions in both groups fall within a similar range, indicating that most items have similar measurement characteristics. However, several items show a larger position gap between the two groups, particularly P1, P5, P14, and P17. This pattern aligns with the DIF statistical analysis, which indicates that these four items have probability values < 0.05. Although there are differences in item positions on the graph, the gaps between the groups remain relatively small; thus, overall, the MTQ-18 demonstrates fairly stable measurement function based on gender.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Table 9. Results of the Pairwise Differential Item Functioning (DIF) Analysis by Gender

Item	DIF Measure Laki-laki	S.E	DIF Measure Perempuan	S.E	DIF Contrast	S.E Contrast	Welch			Mantel-Haenszel		Cumlor Size
							t	d.f	Prob	Chi-Squ	Prob	
1	-0.87	0.11	-0.57	0.07	-0.30	0.13	-2.27	285	0.0242	8.5496	0.0035	-0.63
2	1.35	0.11	1.48	0.07	-0.13	0.13	-0.95	292	0.3421	0.1522	0.6965	-0.08
3	0.29	0.11	0.07	0.07	0.22	0.13	1.69	285	0.0914	3.3430	0.0675	0.41
4	-0.84	0.11	-0.96	0.07	0.11	0.13	0.86	285	0.3927	0.0020	0.9644	0.01
5	-0.43	0.11	-0.82	0.07	0.39	0.13	2.96	285	0.0033	8.5786	0.0034	0.62
6	0.99	0.11	1.08	0.07	-0.09	0.13	-0.66	288	0.5072	0.0778	0.7803	-0.06
7	-0.82	0.11	-0.78	0.07	-0.04	0.13	-0.30	285	0.7618	0.1093	0.7410	-0.07
8	1.57	0.11	1.81	0.08	-0.24	0.14	-1.73	298	0.0853	0.9551	0.3284	-0.21
9	0.69	0.11	0.49	0.07	0.19	0.13	1.46	286	0.1456	4.5541	0.0328	0.45
10	-0.01	0.11	0.09	0.07	-0.10	0.13	-0.73	285	0.4639	1.4898	0.2222	-0.26
11	0.78	0.11	0.74	0.07	0.05	0.13	0.35	286	0.7249	0.6531	0.4190	-0.17
12	0.89	0.11	1.06	0.07	-0.17	0.13	-1.27	288	0.2067	0.4813	0.4878	-0.15
13	-0.83	0.11	-0.94	0.07	0.11	0.13	0.84	285	0.4004	0.0019	0.9649	-0.01
14	-1.61	0.12	-1.94	0.07	0.33	0.14	2.43	287	0.0158	4.9011	0.0268	0.46
15	-1.29	0.11	-1.29	0.07	0.00	0.13	0.00	284	1.0000	0.2795	0.5970	-0.11
16	0.24	0.11	0.24	0.07	0.00	0.13	0.00	285	1.0000	0.1726	0.6778	0.09
17	0.99	0.11	1.34	0.07	-0.35	0.13	-2.63	292	0.0090	2.5026	0.1137	-0.34
18	-1.09	0.11	-1.07	0.07	-0.02	0.13	-0.17	285	0.8652	0.1748	0.6759	-0.08

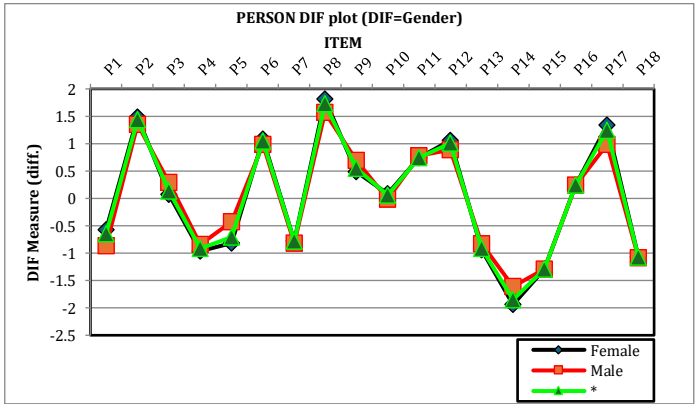


Figure 2. Person DIF Plot (DIF=Gender)

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Overall, the results of the DIF analysis indicate that the MTQ-18 exhibits a fairly good level of measurement equivalence based on gender. The majority of items (14 out of 18) showed no indication of DIF; thus, they can be considered to have relatively equivalent measurement functions between male and female students. Meanwhile, the four items showing indications of DIF can still be retained because the differences in item function found are relatively small and must be considered in conjunction with theoretical fit and the substance of the construct being measured.

3.6. Instrument Reliability

Instrument reliability testing was conducted to assess measurement consistency and the instrument's ability to distinguish between respondent characteristics and item difficulty levels based on the Rasch Model (Bakar et al., 2023). The reliability analysis was based on summary statistics, focusing on three key indicators: Cronbach's Alpha, person reliability, and item reliability (Bakar et al., 2023). The Cronbach's Alpha value is used to describe the internal consistency of the instrument, with the following categories: very poor if the value is less than 0.50; poor in the range of 0.50 to less than 0.60; adequate in the range of 0.60 to less than 0.70; good in the range of 0.70 to 0.80; and very good if greater than 0.80. Meanwhile, person reliability indicates the consistency of estimates of respondents' abilities or characteristics, while item reliability indicates the stability of the item difficulty hierarchy. Referring to the criteria used by Muqaffi et al., (2024), both reliability indicators are categorized as weak if their values are less than 0.67; adequate in the range of 0.67–0.80; good in the range of 0.81–0.90; very good in the range of 0.91–0.94; and exceptional if their values reach 0.95 or higher. Based on these criteria, the results of the instrument reliability analysis using summary statistics are presented in Table 10.

Table 10. Results of the Summary Statistics Analysis

Summary of 455 Measured Person										
	Total Score	Count		Measure		Model Error	Infit		Outfit	
							MNSQ	ZSTD	MNSQ	ZSTD
Mean	52.1	18.0		-.17		.30	1.00	-.2	.99	-.2
SD.	6.5	.0		.58		.01	.59	1.7	.59	1.7
Max.	78.0	18.0		2.37		.39	4.69	6.6	4.77	6.7
Min.	27.0	18.0		-2.79		.29	.19	-3.9	.19	-3.9
Real	RMSE	.33	True	SD	.48	Separation	1.45	Person	Reliability	.68
Model	RMSE	.30	True	SD	.50	Separation	1.67	Person	Reliability	.74
S.E. OF Person Mean = .03										
Person Raw Score-To-Measure Correlation = 1.00										
Cronbach Alpha (KR-20) Person Raw Score "Test" Reliability = .72										

Summary of 455 Measured Item										
	Total Score	Count		Measure		Model Error	Infit		Outfit	
							MNSQ	ZSTD	MNSQ	ZSTD
Mean	1317.9	455.0		.00		.06	.99	-.2	.99	-.3
SD.	297.2	.0		1.03		.00	.17	2.7	.17	2.7
Max.	1847.0	455.0		1.73		.06	1.37	5.4	1.37	5.4
Min.	827.0	455.0		-1.85		.06	.64	-6.5	.64	-6.5
Real	RMSE	.06	True	SD	1.03	Separation	16.76	Item	Reliability	1.00
Model	RMSE	.06	True	SD	1.03	Separation	17.32	Item	Reliability	1.00
S.E. OF Item Mean = .25										
UMEAN = .0000 USCALE = 1.0000										
Item RAW SCORE -TO-MEASURE CORRELATION = -1.00										

The results of the summary statistics analysis showed a Cronbach's Alpha/KR-20 value of 0.72, which falls into the "good" category. This indicates that the instrument has good internal consistency in measuring the construct under study. Furthermore, the person reliability value of 0.68 falls into the "adequate" category, meaning the instrument is sufficiently capable of distinguishing respondents based on their level of ability or tendency toward the construct being measured. Meanwhile, the item reliability value of 1.00 falls into the "excellent" category, indicating that the items in the instrument possess a very high level of consistency and stability. Thus, overall, the mental toughness scale can be deemed reliable, particularly in terms of item quality, although the instrument's ability to distinguish between respondents remains in the "adequate" category.

Overall, the results of the analysis using the Rasch Model indicate that the mental toughness scale possesses good psychometric quality. The instrument meets the assumption of unidimensionality, features varying levels of item difficulty, and demonstrates adequate item fit based on the item fit analysis. Additionally, the response categories on the Likert scale were found to function well and were consistently understood by respondents. The reliability analysis results indicate that the instrument has good internal consistency with very high item quality. Thus, the developed mental toughness scale is deemed valid and can be used as a measurement tool to identify the level of mental toughness among high school students in general.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Mental toughness is one of the key psychological aspects that helps individuals adaptively cope with life's pressures, challenges, and demands (Strycharczyk et al., 2021). For high school students, mental toughness is not only related to the ability to cope with academic pressure but is also linked to the ability to manage emotions, maintain motivation, build self-confidence, and adapt to social changes and personal development (Yusri et al., 2026). Therefore, an instrument capable of accurately measuring mental toughness is essential for supporting educational services and students' psychological development.

The results of this study indicate that the mental toughness scale adapted from the MTQ-18 possesses adequate psychometric characteristics for use with high school students. These findings suggest that instrument adaptation requires not only language translation but also contextual adjustments to ensure that items align with the developmental characteristics of the respondents (Ambuehl & Inauen, 2022). This is important because previous research on mental toughness has largely focused on athletic populations, whereas high school students face different forms of pressure, such as academic demands, social relationships, and the development of self-identity (Yusri et al., 2026). Unlike the study by Dagnall et al., (2019), which developed the MTQ-18 in an educational context with different cultural characteristics, this study provides additional evidence regarding the instrument's functionality among Indonesian high school students. Furthermore, the study by Yudiarso et al., (2025), which tested the Indonesian version of the MTQ-18 using the Rasch approach, focused primarily on an athlete population; thus, this study expands the psychometric analysis of the MTQ-18 to a group of students with distinct characteristics of psychological stress.

The results of the response category analysis indicate that students were able to consistently distinguish between response options on the Likert scale. Increases in the observed average and Andrich threshold values indicate that the response categories on the MTQ-18 function sequentially and are understandable to respondents. These findings align with those of Stefana et al., (2025) who noted that response categories on psychological instruments for adolescents function optimally when each answer option has a clear distinction in meaning. Furthermore, Colledani, Boragno, Fiabane, Setti, and Gabanelli, (2025) explain that well-functioning response categories indicate that the scores obtained are capable of representing the variation in the construct being measured. Thus, the use of a five-point Likert scale in the MTQ-18 is appropriate for measuring high school students' mental toughness.

In terms of item fit, all items remain within the acceptable range of the Rasch model based on the MNSQ Outfit indicator. These findings indicate that, in general, the MTQ-18 items are capable of measuring the construct of mental toughness among high school students. However, further analysis revealed that several items exhibit suboptimal response patterns based on the Outfit ZSTD and Point Measure Correlation indicators, namely P11, P5, P18, P9, P8, P14, P1, P15, P10, and P13. This situation indicates that although the items can still be retained based on Rasch criteria, some items do not yet have a fully strong relationship with the construct being measured. This finding does not directly imply that these items are unsuitable for use, but it suggests that certain aspects of mental toughness may have varying levels of relevance when applied to high school students compared to the original population for which the instrument was developed. This is consistent with the findings of Ambuehl and Inauen, (2022); Cruchinho et al., (2024) who noted that the process of adapting an instrument to a new population can result in variations in item characteristics due to differences in respondents' experiences and interpretations. Instrument evaluation in a new population must consider not only the statistical indicators of the items but also the alignment of the items' meanings with the characteristics of the target group (Syvertsen, Scales, & Toomey, 2021).

The emergence of several items with suboptimal measurement characteristics is likely related to differences in the context of MTQ-18 use. This instrument was originally developed in a sports context that emphasizes an individual's ability to maintain performance when facing competitive pressure. For athletes, pressure is more closely related to competition, achievement, and performance evaluation (Tossici, Zurloni, & Nitri, 2024), whereas high school students face more complex pressures, such as academic demands, social relationships, anxiety about the future, and the process of self-identity formation (Nirmalan, Nidhi, Sunil, & Kachhi, 2025). These differences in experience mean that certain items may be interpreted differently by the two groups. Therefore, adapting the MTQ-18 for high school students not only involves maintaining the instrument's theoretical structure but also ensuring that each item holds psychological relevance to the experiences of adolescents.

The research results also indicate that the MTQ-18 adequately represents the construct of mental toughness among high school students. These findings reinforce the notion that mental toughness is not only relevant in the context of sports but also plays a crucial role in education (Clair-thompson & Devine, 2023). Until now, studies on mental toughness have primarily focused on athletes; however, this construct is also related to students' ability to cope with academic pressure, maintain motivation to learn, and adapt to various developmental demands (Zhang & Khan, 2025). Nevertheless, the results regarding unidimensionality must be interpreted with caution. The raw variance explained by measures, at 47.2%, indicates that the instrument measures the same primary construct; however, theoretically, mental toughness still comprises several interrelated aspects: control, commitment, challenge, and confidence. Thus, the fulfillment of the

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

unidimensionality assumption in the Rasch model does not mean that the construct is entirely singular but rather indicates that these various dimensions have strong relationships in forming the general construct of mental toughness. Therefore, further research is still needed to test whether this four-dimensional structure remains consistent among Indonesian high school students through factor analysis or other measurement models.

Variations in item difficulty levels indicate that the MTQ-18 is capable of assessing students with different levels of mental toughness. More difficult items generally relate to the ability to manage stress, control negative thoughts, and maintain focus when facing challenging situations. These abilities are still developing during adolescence because students are in a stage of emotional and social development (Suryana, Tika, & Wardani, 2022). Conversely, easier items relate to self-confidence and optimism, which are more closely tied to students' everyday experiences. These findings indicate that mental toughness in adolescents is a psychological ability that develops through experiences of facing various life challenges (Arslan & Wong, 2024).

Regarding reliability, the study results show that the stability of the item hierarchy falls into the "very good" category, but the instrument's ability to distinguish between respondents' characteristics remains in the "adequate" category. The difference between high item reliability and lower person reliability indicates that the item hierarchy is stable, but variations in mental toughness levels among respondents have not yet been fully and optimally distinguished. This condition is likely influenced by the relatively homogeneous sample characteristics, as the participants came from the same educational level and age range (Franzen et al., 2022). In the Rasch Model, respondent reliability is influenced by the instrument's ability to capture variations in individual ability across the range of the construct being measured (Bakar et al., 2023). Thus, low person reliability does not merely indicate a weakness in the instrument but may also reflect limitations in the variation of respondent characteristics in this study. Therefore, future research should involve a more heterogeneous sample so that the instrument's ability to distinguish levels of individual mental toughness can be more robustly tested.

In addition to reliability, the results of the DIF analysis indicate that most MTQ-18 items have relatively equivalent measurement functions among male and female students. However, there are several items that show indications of DIF, suggesting that certain items may be interpreted differently based on gender. These findings suggest that instrument validation is not only about score consistency but also about ensuring that items measure the same construct fairly across different respondent groups. This aligns with Graves, Hall, Dias-Karch, Haischer, and Apter, (2021), who noted that differences in social experiences, emotional regulation, and coping strategies under pressure can be factors causing variations in responses between male and female students. Therefore, items showing indications of DIF need to be further examined through follow-up research to ensure that score differences truly reflect variations in mental toughness, rather than resulting from differences in how the items are understood.

Overall, this study indicates that the adapted MTQ-18 mental toughness scale provides preliminary evidence of adequate psychometric quality and is relevant for use with high school students. Compared to previous studies that have primarily focused on athletes, this research contributes by expanding the application of the MTQ-18 to an educational context. However, several aspects still need to be developed, particularly the evaluation of specific items, improving the instrument's ability to distinguish between respondents, testing its multidimensional structure, and evaluating item equivalence based on student characteristics. Thus, the MTQ-18 has the potential to be used as an initial assessment instrument in the fields of education and guidance counseling; however, the interpretation of results must still take into account the characteristics of the population and the measurement context.

3.7. Implications

The results of this study have implications for the development of Guidance and Counseling (BK) services in schools, particularly in the process of assessing students' needs. The adapted MTQ-18 shows preliminary evidence of adequate psychometric quality, making it a viable instrument for assessing high school students' levels of mental toughness. This information can help guidance and counseling teachers understand students' abilities to cope with stress, maintain motivation, manage emotions, and adapt to developmental demands. In the practice of guidance and counseling services, mental toughness assessment results can serve as the basis for developing need-based programs. In basic counseling services, assessment data can be used to design individual or group counseling sessions focused on strengthening psychological resilience, regulating emotions, and developing the ability to cope with academic pressure. In responsive counseling services, assessment results can help identify students who require further support through individual counseling, group counseling, or preventive interventions. Furthermore, mental toughness data can support individual planning and systemic support through student self-development, educational and career planning, evidence-based evaluation of guidance and counseling programs, and collaboration with schools and families. However, the MTQ-18 must be used with caution, as this instrument is more appropriately used as an initial assessment rather than as a tool for psychological diagnosis. The study's limitations regarding sample characteristics, the instrument's ability to distinguish between respondents, and the evaluation of certain items highlight the need for further research

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

with a more diverse population to strengthen the empirical basis for using the MTQ-18 in the Indonesian educational context.

3.8. Limitations

The limitations of this study relate to the relatively limited and homogeneous sample characteristics, as all participants were high school students in Sidrap Regency with similar age ranges and educational levels. These conditions mean that the study's results must be interpreted with caution when applied to a population of high school students with more diverse cultural, social, and academic backgrounds. Furthermore, future research should involve a broader and more heterogeneous sample to strengthen the evidence supporting the generalizability of the instrument. Although the Rasch analysis results indicate adequate psychometric characteristics, further testing regarding construct structure, measurement consistency across different groups, and the development of validity evidence for the MTQ-18 in the context of Indonesian education are still needed.

4. Conclusion

This study shows that the Mental Toughness Questionnaire-18 (MTQ-18), as adapted for high school students, provides preliminary evidence of adequate psychometric quality based on Rasch Model analysis. This instrument demonstrates good response category functionality, item fit with the measurement model, fulfillment of the unidimensionality assumption, variation in item difficulty levels, and good internal consistency with very high item quality. Although test-retest reliability is in the "adequate" category, the MTQ-18 still shows potential as an initial assessment tool (need assessment) to help guidance and counseling teachers understand students' levels of mental toughness and develop needs-based services. This study expands the evidence for the use of the MTQ-18 in an educational context, particularly among non-athlete high school students. However, the generalizability of the findings is still limited to the characteristics of the study sample; therefore, future research should involve a broader and more heterogeneous sample and conduct further testing of the construct structure and item functional equivalence to strengthen the evidence of the MTQ-18's validity in the Indonesian educational context.

Author Contributions

All authors made equal contributions to the preparation and completion of this article. All authors have read, reviewed, and approved the final version of the manuscript for publication.

Funding

This research received funding support from the Institute for Research and Community Service (LPPM) at Malang State University for Fiscal Year 2026, as outlined in Research Implementation Agreement No. 14.04.1057/UN32.14.1/LT/2026.

Declaration of Conflicting Interests

The authors declare that there are no conflicts of interest in the research, the writing process, or the publication of this article. This study has received ethical approval from the Ethics Committee of Malang State University, with approval number 16.07.24/UN32.14.2.8/LT/2026. All research procedures were conducted in accordance with applicable research ethics principles, including the provision of information regarding the research objectives, data collection procedures, the confidentiality of respondent information, and voluntary participation.

Ethics and Consent Declarations

This study has received ethical approval from the Ethics Committee of Malang State University via approval letter No. 16.07.24/UN32.14.2.8/LT/2026. All stages of the study were conducted in accordance with applicable research ethics principles and regulations. This study is also officially registered with the Research Institute of Malang State University under Protocol No. 01492335731112520260617001 71. Prior to data collection, participants were informed about the research objectives, the procedures for completing the instruments, and assurances regarding data confidentiality. The researchers also emphasized that participation in this study was voluntary, so participants were free to agree or decline to participate without any consequences. Participants were also informed that all data collected would be used solely for research purposes and analyzed without disclosing their personal identities.

Data Availability

Data used and/or analyzed in this study are available from the corresponding author upon reasonable request.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Declaration on AI Use

The authors declare that ChatGPT, developed by OpenAI, was used to a limited extent in this study to help explore alternative wording for several instrument items to make them simpler and easier for high school students to understand, as well as to help improve the readability and linguistic quality of the manuscript. The use of AI was limited to editorial aspects and was not used to determine constructs, indicators, instrument validity, data analysis, interpretation of results, or drawing conclusions. All generated wording alternatives were reviewed by the authors, while the content equivalence of the items after adjustment was verified by a guidance and counseling expert. All scientific decisions, analyses, interpretations, and conclusions in this article are the sole responsibility of the authors.

References

Abdellatif, H. (2023). Test results with and without blueprinting: Psychometric analysis using the Rasch model. *Educación Médica, 24*(3), Article 100802. https://doi.org/10.1016/j.edumed.2023.100802

Abu Bakar, S. N., Mahmood, N., Ismail, A., Hashim, A., Hamsan, N. Y., Nor Hamid, N. I., & Che Yakzam, R. (2023). Validity and reliability of competency analysis instrument for cooperative board members using Rasch measurement model approach. *International Journal of Academic Research in Business and Social Sciences, 13*(6), 2408–2419. https://doi.org/10.6007/IJARBSS/v13-i6/17543

Adelina, I., Analya, P., & Anggelia, Y. (2022). Peran mindset terhadap ketangguhan mental mahasiswa. *Syntax Literate: Jurnal Ilmiah Indonesia, 7*(9), 16089–16105. https://doi.org/10.36418/syntax-literate.v7i9.13794

Albano, T., French, B. F., & Vo, T. T. (2024). Traditional vs intersectional DIF analysis: Considerations and a comparison using state testing data. *Applied Measurement in Education, 37*(1), 57–70. https://doi.org/10.1080/08957347.2024.2311935

Almaliyah, S., & Ahman. (2024). Validasi instrumen sikap optimisme remaja: Analysis Rasch model. *G-Couns: Jurnal Bimbingan dan Konseling, 9*(1), 701–709. https://doi.org/10.31316/gcouns.v9i1.6310

Ambuehl, B., & Inauen, J. (2022). Contextualized measurement scale adaptation: A 4-step tutorial for health psychology research. *International Journal of Environmental Research and Public Health, 19*(19), Article 12775. https://doi.org/10.3390/ijerph191912775

Amiruddin, M. Z. (2026). Assessing educational instrument validity-reliability and student performance: Rasch model insights. *Asian Journal of Assessment in Teaching and Learning, 16*(1), 12–27. https://doi.org/10.37134/ajatel.vol16.1.2.2026

Anggara, N., Rusmana, N., & Saripah, I. (2025). Adaptation of student interpersonal relationship measurements: Validity and reliability of instruments with the Rasch model. *G-Couns: Jurnal Bimbingan dan Konseling, 10*(1), 766–776. https://doi.org/10.31316/g-couns.v10i01.6272

Arisman. (2024). Faktor faktor yang mempengaruhi mental toughness pada atlit cabang olahraga permainan. *Prosiding Seminar Nasional Pendidikan Jasmani dan Kesehatan Mental Peserta Didik*, 645–662. https://proceedings.unnes.ac.id/snpjkmpd/article/view/4237

Arslan, G., & Wong, P. (2024). Embracing life's challenges: Developing a tool for assessing resilient mindset in second wave positive psychology. *Journal of Happiness and Health, 4*(1), 1–10. https://doi.org/10.47602/johah.v4i1.53

Asradi, & 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*(3), 2406–2419. https://doi.org/10.31316/g-couns.v10i03.8722

Bandura, A. (2023). Cultivate self-efficacy for personal and organizational effectiveness. In E. Locke & C. Pearce (Eds.), *Principles of organizational behavior: The handbook of evidence-based management* (3rd ed.). Wiley. https://doi.org/10.1002/9781394320769.ch6

Boonma, M. (2025). The power of resilience in 21st-century education: How growth mindset and mental well-being support teachers and students. *Nakhon Ratchasima Journal of Humanities and Social Sciences, 1*(3), 1–26.

Chiu, T. K. F., Chen, Y., Yau, K. W., Chai, C., Meng, H., King, I., Wong, S., & Yam, Y. (2024). Developing and validating mea sures for AI literacy tests: From self-reported to objective measures. *Computers and Education: Artificial Intelligence, 7*, Article 100282. https://doi.org/10.1016/j.caeai.2024.100282

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. *Jurnal Cakrawala Pendidikan, 41*(1), 97–111. https://doi.org/10.21831/cp.v41i1.39130

Cioffi, A., Galano, S., Passeggia, R., & Testa, I. (2024). Validation of two test anxiety scales for physics undergraduate co urses through confirmatory factor analysis and Rasch analysis. *Physical Review Physics Education Research, 20*(1), Article 010126. https://doi.org/10.1103/PhysRevPhysEducRes.20.010126

Clough, P., Earle, K., & Sewell, D. (2002). Mental toughness: The concept and its measurement. In I. Cockerill (Ed.), *Solutions in sport psychology* (pp. 32–46). Thomson.

Colledani, D., Boragno, P., Fiabane, E. M., Setti, I., & Gabanelli, P. (2025). Further validation of the Persecutory Ideation Questionnaire in the Italian context: Results from classical and modern test theory. *Assessment, 32*(4), 577–589. https://doi.org/10.1177/10731911241257012

Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Miranda da Silva, M., Teixeira, G., Nunes, E., Lucas, P., & Gaspar, F. (2024). Translation, cross-cultural adaptation, and validation of measurement instruments: A practical

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

guideline for novice researchers. *Journal of Multidisciplinary Healthcare, 17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714

Dagnall, N., Denovan, A., Papageorgiou, K. A., Clough, P. J., Parker, A., & Drinkwater, K. G. (2019). Psychometric assessment of shortened Mental Toughness Questionnaires (MTQ): Factor structure of the MTQ-18 and the MTQ-10. *Frontiers in Psychology, 10*, Article 1933. https://doi.org/10.3389/fpsyg.2019.01933

Darmawan, A. A., Zamzir, Z., Widiasiti, A. A. I. P., Sugiharni, G. A. D., & Amir, F. L. (2025). *Pengembangan instrumen penelitian*. Naga Pustaka.

Deng, Y., Cherian, J., Khan, N. U. N., Kumari, K., Sial, M. S., Comite, U., Gavurova, B., & Popp, J. (2022). Family and academic stress and their impact on students' depression level and academic performance. *Frontiers in Psychiatry, 13*, Article 869337. https://doi.org/10.3389/fpsyt.2022.869337

Denovan, A., Dagnall, N., Artamonova, E., & Papageorgiou, K. A. (2024). Evaluation of the 10-item Mental Toughness Questionnaire (MTQ10): Cross-cultural assessment and scrutiny of method effects. *Current Psychology, 43*(40), 31254–31266. https://doi.org/10.1007/s12144-024-06562-9

Diba, S. F., & Ahman. (2024). Validitas inventori resiliensi akademik siswa sekolah menengah atas dengan pendekatan Rasch model. *G-Couns: Jurnal Bimbingan dan Konseling, 9*(1), 619–627. https://doi.org/10.31316/gcouns.v9i1.6354

Engelhard, G., Jr., & Wang, J. (2024). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences* (2nd ed.). Routledge.

Fadli, R., Hidayati, S., Cholifah, M., Siroj, R. A., & Afgani, M. W. (2023). Validitas dan reliabilitas pada penelitian motivasi belajar pendidikan agama Islam menggunakan product moment. *JIIP: Jurnal Ilmiah Ilmu Pendidikan, 6*(3), 1734–1739. https://doi.org/10.54371/jiip.v6i3.1419

Fajar, A. S. M., Zakki, M., Darmawan, D., Evendi, W., & Mujito. (2021). A socio-psychological analysis of adolescents' choice of ideals in an environment of social pressure. *Journal of Social Science Studies, 1*(1), 251–256.

Farnsworth, J. L., Marshal, A., & Myers, N. L. (2022). Mental toughness measures: A systematic review of measurement properties for practitioners. *Journal of Applied Sport Psychology, 34*(3), 479–494. https://doi.org/10.1080/10413200.2020.1866710

Fengjun, Q., Jing, Z., & Liguang, L. (2022). The effect of mental toughness on learning burnout of junior middle school students: Putting school adaptation as a mediator variable. *Discrete Dynamics in Nature and Society, 2022*(1), Article 9706046. https://doi.org/10.1155/2022/9706046

Fernandes, A. A. R., & Akhrani, L. A. (2022). *Rancangan pengukuran variabel: Angket dan kuesioner (pemanfaatan R)*. Universitas Brawijaya Press.

Flores-Builes, R., & Mateu-Pérez, R. (2025). Influence of decision-making styles on academic stress and resilience: An approach for educational intervention. *Psychology in the Schools, 62*(10), 4170–4179. https://doi.org/10.1002/pits.23603

Folastri, S., Hambali, I., Ramli, M., Akbar, S., & Sofyan, A. (2023). Development and preliminary validation of a scale to measure sexual violence awareness using the Rasch model. *Psychology Hub, 40*(3), 59–65. https://doi.org/10.13133/2724-2943/18000

Franzen, S., Smith, J. E., van den Berg, E., Rivera Mindt, M., van Bruchem-Visser, R. L., Abner, E. L., Schneider, L. S., Prins, N. D., Babulal, G. M., & Papma, J. M. (2022). Diversity in Alzheimer's disease drug trials: The importance of eligibility criteria. *Alzheimer's & Dementia, 18*(4), 810–823. https://doi.org/10.1002/alz.12433

Golzar, J., Noor, S., & Tajik, O. (2022). Convenience sampling. *International Journal of Education & Language Studies, 1*(2), 72–77. https://doi.org/10.22034/ijels.2022.162981

Graves, B. S., Hall, M. E., Dias-Karch, C., Haischer, M. H., & Apter, C. (2021). Gender differences in perceived stress and coping among college students. *PLOS ONE, 16*(8), Article e0255634. https://doi.org/10.1371/journal.pone.0255634

Guszkowska, M., & Wójcik, K. (2021). Effect of mental toughness on sporting performance: Review of studies. *Baltic Journal of Health and Physical Activity, 13*(Suppl. 2), 1–12. https://doi.org/10.29359/BJHPA.2021.Suppl.2.01

Hadžibajramović, E., Schaufeli, W., & De Witte, H. (2022). Shortening of the Burnout Assessment Tool (BAT)—From 23 to 12 items using content and Rasch analysis. *BMC Public Health, 22*, Article 560. https://doi.org/10.1186/s12889-022-12946-y

Hafiyani, S., & Ahman. (2024). Pengembangan dan validasi instrumen skala nonviolence siswa sekolah menengah kejuruan (SMK) di Kota Bandung: Analisis pemodelan Rasch. *G-Couns: Jurnal Bimbingan dan Konseling, 9*(1), 74–88. https://doi.org/10.31316/gcouns.v9i1.6230

Haqiyah, A., Sanjaya, K. H., Soeharto, Tangkudung, A. W. A., Riyadi, D. N., Lubis, J., Pratiwi, E., Puspitorini, W., Mulyadi, D. Y. N., & Kusumandari, D. E. (2023). Developing and validating the mental toughness questionnaire of athletes using Rasch analysis. *International Journal of Human Movement and Sports Sciences, 11*(3), 650–659. https://doi.org/10.13189/saj.2023.110319

Hsieh, Y.-C., Lu, F. J. H., Gill, D. L., Hsu, Y.-W., Wong, T.-L., & Kuan, G. (2024). Effects of mental toughness on athletic performance: A systematic review and meta-analysis. *International Journal of Sport and Exercise Psychology, 22*(6), 1317–1338. https://doi.org/10.1080/1612197X.2023.2204312

Ilfiandra, Nadhirah, N. A., Suryana, D., & Ahmad, A. B. (2022). Development and validation peaceful classroom scale: Rasch model analysis. *International Journal of Instruction, 15*(4), 497–514. https://doi.org/10.29333/iji.2022.15427a

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Ismail, Riandi, Kaniawati, I., Sopandi, W., Supriyadi, Suhendar, & Hidayat, F. A. (2024). Gender roles in understanding and implementing green energy technology in Indonesian schools: Rasch analysis. *Qubahan Academic Journal, 4*(3), 298–314. https://doi.org/10.48161/qaj.v4n3a752

Janjani, P., Nosrati, P., Ziapour, A., Khezeli, M., Geravand, S., Heidari Moghadam, R., & Yıldırım, M. (2025). Mediating role of coping strategies in the relationship between disease management self-efficacy, mental toughness, and quality of life in cardiovascular patients. *Discover Mental Health, 5*, Article 22. https://doi.org/10.1007/s44192-025-00133-5

Jiang, Y., Fu, Y., & Dong, X. (2025). Effects of physical exercise on college students' academic self-efficacy: The chain mediating role of future orientation and mental toughness. *Frontiers in Psychology, 16*, Article 1604725. https://doi.org/10.3389/fpsyg.2025.1604725

Kamila, A., & Widodo, P. B. (2025). Validitas konstruk dalam psikometri: Peran analisis faktor eksploratori dan konfirmatori. *Jurnal Psikologi Pendidikan dan Pengembangan SDM, 13*(1), 1–7.

Kiarostami, A., Jeddi, M., Roohinezhad, S., & Hakimi, R. (2022). The effectiveness of positive thinking training on mental toughness and social well-being in derelict adolescents. *Iranian Evolutionary and Educational Psychology Journal, 4*(2), 398–409. https://doi.org/10.52547/ieepj.4.2.398

Latifah, M., Saripah, I., Suryana, D., & Sunarya, Y. (2024). Validity and reliability of self-concept instrument using Rasch model. *Jurnal Kajian Bimbingan dan Konseling, 9*(1), 26–35. https://doi.org/10.17977/um001v9i12024p26-35

Lee, K., & Lee, H. (2025). Comparative study through item analysis of CCRAM28 and CCRAM10. *Social Indicators Research, 177*(1), 411–424. https://doi.org/10.1007/s11205-025-03567-0

Lee, M., & Kim, B. (2023). Effect of the employees' mental toughness on organizational commitment and job satisfaction: Mediating psychological well-being. *Administrative Sciences, 13*(5), Article 133. https://doi.org/10.3390/admsci13050133

Maier, C., Thatcher, J. B., Grover, V., & Dwivedi, Y. K. (2023). Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. *International Journal of Information Management, 70*, Article 102625. https://doi.org/10.1016/j.ijinfomgt.2023.102625

Mansur, D., Aditama, M. H. R., Estafianto, H. D., & Laras, P. B. (2026). Validation of the Indonesian version of the Student-Teacher Relationship Scale using Rasch analysis. *G-Couns: Jurnal Bimbingan dan Konseling, 10*(3), 2520–2540. https://doi.org/10.31316/g-couns.v10i03.9351

Mardhiyah, Dinilhaq, N. A., Amelia, Y., Arini, A., Hidayatullah, R., & Harmonedi. (2025). Populasi dan sampel dalam penelitian pendidikan: Memahami perbedaan, implikasi, dan strategi pemilihan yang tepat. *Katalis Pendidikan: Jurnal Ilmu Pendidikan dan Matematika, 2*(2), 208–218. https://doi.org/10.62383/katalis.v2i2.1670

Maulidya, N., & Jannah, M. (2021). Hubungan psychological capital dengan mental toughness pada atlet beladiri. *Character: Jurnal Penelitian Psikologi, 8*(6), 106–116. https://doi.org/10.26740/cjpp.v8i6.41967

Medvedev, O. N., & Krägeloh, C. U. (2022). Rasch measurement model. In O. N. Medvedev, C. U. Krägeloh, R. J. Siegert, & N. N. Singh (Eds.), *Handbook of assessment in mindfulness research* (pp. 1–18). Springer. https://doi.org/10.1007/978-3-030-77644-2_4-1

Muqaffi, Rusmana, N., & Ilfiandra. (2024). Emotional regulation scale: A Rasch model approach. *G-Couns: Jurnal Bimbingan dan Konseling, 9*(1), 451–459. https://doi.org/10.31316/gcouns.v9i1.6453

Muslihin, H. Y., Suryana, D., Ahman, Suherman, U., & Dahlan, T. H. (2022). Analysis of the reliability and validity of the self-determination questionnaire using Rasch model. *International Journal of Instruction, 15*(2), 207–222. https://doi.org/10.29333/iji.2022.15212a

Nirmalan, A., Nidhi, N., Sunil, R., & Kachhi, Z. (2025). Effects of the high school to college transition on student well-being, personal growth, and resilience building. In *Nurturing student well-being in the modern world* (pp. 171–194). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-1455-5.ch007

Norhayati, M. N., Fatin Imtithal, A., & Nor Akma, Y. (2021). Psychometric properties of the Malay version Women's Views of Birth Postnatal Satisfaction Questionnaire using the Rasch measurement model: A cross-sectional study. *BMC Pregnancy and Childbirth, 21*, Article 711. https://doi.org/10.1186/s12884-021-04184-8

Pahabol, N., Guntoro, T. S., Sutoro, Fariz, M., & Putra, P. (2024). Bagaimana gambaran ketangguhan mental atlet? Sebuah studi deskriptif. *Multilateral: Jurnal Pendidikan Jasmani dan Olahraga, 23*(4), 141–148. https://doi.org/10.20527/multilateral.v23i4.20963

Perry, J. L., Strycharczyk, D., Dagnall, N., Denovan, A., Papageorgiou, K. A., & Clough, P. J. (2021). Dimensionality of the Mental Toughness Questionnaire (MTQ48). *Frontiers in Psychology, 12*, Article 654836. https://doi.org/10.3389/fpsyg.2021.654836

Pratama, E. Y., & Utami, R. J. (2024). *Psikologi olahraga: Pendekatan holistik bagi atlet*. Pradina Pustaka.

Prayoga, K. P., Suryana, D., Supriatna, M., & Budiman, N. (2024). Penggunaan Rasch model untuk menganalisis konstruk instrumen kontrol diri pada siswa sekolah menengah. *G-Couns: Jurnal Bimbingan dan Konseling, 9*(1), 367–381. https://doi.org/10.31316/gcouns.v9i1.4459

Putri, E. K., & Febriani, A. (2021). Academic self-efficacy as the mediator of adolescent-parent secure attachment effect towards the academic stress of senior high school students. *Jurnal Psikologi, 48*(2), 173–182. https://doi.org/10.22146/jpsi.52491

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Putri, N. L. M. W. D., Suranata, K., & Lestari, L. P. (2021). Rasch career self efficacy scale analysis for high school students. *Bisma: The Journal of Counseling, 5*(3), 204–212. https://doi.org/10.23887/bisma.v5i3.42031

Rahma, U. H., Azmi, M. M., Utama, M. E., & Triastuti, E. P. R. (2025). Memperkuat mental toughness melalui appreciative inquiry pada atlet pencak silat Persinas ASAD. *Motiva: Jurnal Psikologi, 8*(1), 61–73.

Rahmania, S., & Royanto, L. R. (2021). Adaptasi alat ukur keterlibatan pembelajar daring pada mahasiswa di Indonesia. *Edcomtech: Jurnal Kajian Teknologi Pendidikan, 6*(2), 173–185. https://doi.org/10.17977/um039v6i12021p173

Ramadani, U. P., Muthmainnah, R., Ulhilma, N., Wazabirah, A., Hidayatullah, R., & Harmonedi. (2025). Strategi penentuan populasi dan sampel dalam penelitian pendidikan: Antara validitas dan representativitas. *QOSIM: Jurnal Pendidikan, Sosial & Humaniora, 3*(2), 574–585. https://doi.org/10.61104/jq.v3i2.1021

Rumtini, Fauzan, A., Suwondo, T., Putra, I. K. S., Tama, I. W., Herlinawati, Sundari, K. N., Astuti, M. T., Rema, I. N., Setiawan, B., Mardiyono, & Pamungkas, O. Y. (2026). Compliance behavior with Indonesian language in local government administration: An analysis of the Rasch model. *Research Journal in Advanced Humanities, 7*(1), 1–19. https://doi.org/10.58256/f7ynsr76

Sari, E. D. K., & Mahmudi, I. (2019). *Analisis pemodelan Rasch pada asesmen pendidikan (analisis dengan menggunakan aplikasi Winstep)*. https://www.researchgate.net/profile/Eva-Sari-3/publication/378210331_BUKU_ANALISIS_PEMODELAN_RASCH_PADA_ASESMEN_PENDIDIKAN/links/65cd79b81bed776ae35f51e9/BUKU-ANALISIS-PEMODELAN-RASCH-PADA-ASESMEN-PENDIDIKAN.pdf

Segara, S. C., Salma, I., & Siregar, P. A. (2025). Menumbuhkan semangat belajar remaja: Kajian psikologi pendidikan tentang faktor internal dan eksternal motivasi. *Journal of Sustainable Education (JOSE), 2*(2), 280–288. https://doi.org/10.63477/jose.v2i2.223

Setiawati, R., Arthur, R., & Khamdan, M. (2024). Analisis model Rasch instrumen tes pemahaman peserta pendidikan pelatihan dasar CPNS petugas pemasyarakatan pada materi agenda dua nilai-nilai dasar BerAKHLAK. *Visipena, 15*(2), 189–203. https://doi.org/10.46244/visipena.v15i2.2989

Setiyowati, A. J., Jaafar, W. M. W., Apriani, R., Prihatiningsih, R., & Ramadhani, E. (2025). Development and validation of the emotional regulation inventory (ERI) using the Rasch model. *Journal of Education and Health Promotion, 14*, Article 488. https://doi.org/10.4103/jehp.jehp_265_25

Singh, H. R., & Bisht, B. (2025). Youth mental health and academic stress: Challenges and pathways for. *Indian Journal of Social Sciences and Literature Studies, 11*(2), 115–122.

St Clair-Thompson, H., & Devine, L. (2023). Mental toughness in higher education: Exploring the roles of flow and feedback. *Educational Psychology, 43*(4), 326–343. https://doi.org/10.1080/01443410.2023.2205622

Stefana, A., Damiani, S., Granziol, U., Provenzani, U., Solmi, M., Youngstrom, E. A., & Fusar-Poli, P. (2025). Psychological, psychiatric, and behavioral sciences measurement scales: Best practice guidelines for their development and validation. *Frontiers in Psychology, 15*, Article 1494261. https://doi.org/10.3389/fpsyg.2024.1494261

Strycharczyk, D., Clough, P., & Perry, J. (2021). *Developing mental toughness: Strategies to improve performance, resilience and wellbeing in individuals and organizations*. Kogan Page.

Subekti, M. R. (2024). Peran mental toughness peserta didik dalam pendidikan jasmani olahraga dan kesehatan di sekolah dasar. *Prosiding Seminar Nasional Pendidikan Jasmani dan Kesehatan Mental Peserta Didik*, 158–174. https://proceedings.unnes.ac.id/snpjkmpd/article/view/4191

Supriatna, M., Suryana, D., Anzhali, M. N., & Noor, A. M. (2025). Evaluating career self-awareness instruments for victims of violence using the Rasch model. *BMC Psychology, 13*, Article 456. https://doi.org/10.1186/s40359-025-02762-5

Suryana, D., Tika, R., & Wardani, E. K. (2022). Management of creative early childhood education environment in increasing golden age creativity. In *Proceedings of the 6th International Conference of Early Childhood Education (ICECE-6 2021)* (pp. 17–20). Atlantis Press. https://doi.org/10.2991/assehr.k.220602.005

Syvertsen, A. K., Scales, P. C., & Toomey, R. B. (2021). Developmental assets framework revisited: Confirmatory analysis and invariance testing to create a new generation of assets measures for applied research. *Applied Developmental Science, 25*(4), 291–306. https://doi.org/10.1080/10888691.2019.1613155

Tanjung, A., Supriatna, M., & Nadhirah, N. A. (2024). Pengembangan instrumen kemandirian remaja menggunakan analisis Rasch model. *Psychocentrum Review, 6*(3), 148–158. https://doi.org/10.26539/pcr.632320

Tennant, A., & Küçükdeveci, A. A. (2023). Application of the Rasch measurement model in rehabilitation research and practice: Early developments, current practice, and future challenges. *Frontiers in Rehabilitation Sciences, 4*, Article 1208670. https://doi.org/10.3389/fresc.2023.1208670

Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2024). Interpreting results from Rasch analysis 1. The "most likely" measures coming from the model. *Disability and Rehabilitation, 46*(3), 591–603. https://doi.org/10.1080/09638288.2023.2169771

Tesio, L., Caronni, A., Simone, A., Kumbhare, D., & Scarano, S. (2024). Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment. *Disability and Rehabilitation, 46*(3), 604–617. https://doi.org/10.1080/09638288.2023.2169772

Tossici, G., Zurloni, V., & Nitri, A. (2024). Stress and sport performance: A PNEI multidisciplinary approach. *Frontiers in Psychology, 15*, Article 1358771. https://doi.org/10.3389/fpsyg.2024.1358771

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Vauziah, V., Permadi, A. A., & Kurniawati, A. (2026). Studi diagnostik profil mental toughness dan kecemasan bertanding atlet petanque BK Porprov Garut 2025. *Jurnal Porkes, 9*(1), 530–543. https://doi.org/10.29408/porkes.v9i1.33635

Yudha, R. P. (2023). Higher order thinking skills (HOTS) test instrument: Validity and reliability analysis with the Rasch model. *EduMa: Mathematics Education Learning and Teaching, 12*(1), 21–38.

Yudiarso, A., Ardhiani, I. W., Surya, R., Watimena, F. Y., & Kanzaki, M. (2025). Psychometric properties of the 18-item Indonesian Mental Toughness Questionnaire using the Rasch model and machine learning. *Psikohumaniora: Jurnal Penelitian Psikologi, 10*(1), 1–20. https://doi.org/10.21580/pjpp.v10i1.23214

Yusri, A. G., Indreswari, H., Bariyyah, K., Setyowati, A. J., & Rahman, D. H. (2026). Bertahan atau menyerah: Gambaran mental toughness siswa SMA dalam menghadapi tekanan. *Guiding World (Bimbingan dan Konseling), 9*(1), 72–84. https://doi.org/10.33627/gw.v9i1.4385

Zhang, S., Li, C., & Unger, D. L. (2022). International doctoral students' sense of belonging, mental toughness, and psychological well-being. *Journal of Comparative & International Higher Education, 14*(2), 138–151. https://doi.org/10.32674/jcihe.v14i1.3432

Zhang, X., & Khan, S. U. (2025). Fostering self-efficacy in Chinese university athletes: The mediating roles of psychological resilience and emotional regulation and the moderating role of autonomy-supportive coaching. *Frontiers in Psychology, 16*, Article 1664339. https://doi.org/10.3389/fpsyg.2025.1664339

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174