Edusainstika : Jurnal Pembelajaran MIPA
Penerbit: Universitas Islam Negeri Mahmud Yunus Batusangkar
Website: https://ejournal.uinmybatusangkar.ac.id/ojs/index.php/Edusainstika
E-mail: edusainstika@uinmybatusangkar.ac.id

Juli 2026
Vol 6 No 1
ISSN: 2807-9388 (e)
pp : 1-18

Comparing Physics Test Anxiety between Religious-Based and Public Senior High School Students Using Rasch Measurement and Differential Item Functioning (DIF)

EN Izzati*[1], MI Sukarelawan[1], M Base[2], RI Ubao[3], S Fayanto[4]

[1]Department of Physics Education, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[2]University of Saint La Salle, Bacolod, Philippines
[3]Adamson University, Manila, Philippines
[4]Institute for Innovative Learning, Mahidol University, Nakhon Pathom, Thailand

elsarina.n@gmail.com

Abstract. Physics test anxiety has been recognized as an important factor affecting students' academic performance; however, evidence regarding differences in physics test anxiety across school contexts remains limited. This study compared physics test anxiety between students from religious-based and public senior high schools and examined whether the Test Anxiety Inventory for Physics (TAIP) functioned equivalently across the two groups using Differential Item Functioning (DIF) analysis within the Rasch measurement framework. A quantitative survey was conducted involving 232 senior high school students from Yogyakarta and Central Java, Indonesia. The Indonesian TAIP was translated through forward–backward translation and administered on a five-point Likert scale. Rasch analysis was employed to estimate person measures, evaluate item fit, and identify DIF, while differences in overall anxiety levels were examined using the Mann–Whitney U test. The findings revealed that students from both school types generally exhibited low levels of physics test anxiety, with no statistically significant difference in overall anxiety levels between religious-based and public senior high schools ($U = 6096.50$, $p = 0.437$, $Z = -0.776$). Nevertheless, the DIF analysis identified three items (WR2, ES5, and LCS3) that functioned differently across school contexts, indicating differences in students' responses to specific aspects of physics test anxiety despite comparable overall anxiety levels. These findings suggest that total anxiety scores may overlook item-level differences. Combining group comparisons with DIF analysis enhances the assessment of physics test anxiety and measurement equivalence across school contexts.

Keywords: Physics Test Anxiety; Rasch Measurement; Differential Item Functioning; Religious-Based Senior High School; Public Senior High School

1. Introduction

Physics is one of the fundamental scientific disciplines underpinning the advancement of various fields of science and modern technology. Major scientific breakthroughs, ranging from the discovery of atomic structure to the development of quantum mechanics, have their origins in fundamental physics research (Bobur et al., 2021). Consequently, physics education plays a strategic role in developing students' scientific reasoning, problem-solving skills, and scientific literacy, which are essential for addressing future technological challenges (Wang, 2023). However, the abstract nature of physics concepts, their mathematical complexity, and the high level of reasoning required often lead secondary school students to perceive physics as one of the most challenging subjects. This complexity not only affects students' conceptual understanding but also influences their psychological readiness when facing various forms of

learning assessment, particularly physics examinations, which serve as one of the primary indicators of academic achievement (Phafiandita & Permadani, 2022).

Test anxiety is closely related to the broader concept of anxiety, which is characterized by feelings of fear, worry, and psychological discomfort when individuals encounter situations perceived as challenging or beyond their ability to control. Such anxiety not only represents an emotional response but also influences students' perceptions of their own competence and academic readiness. Consequently, heightened anxiety may reduce self-confidence, interfere with cognitive processing, and hinder students from demonstrating their actual abilities during examinations. These characteristics highlight that test anxiety should be understood as a multidimensional psychological construct involving emotional, cognitive, and behavioral responses that may ultimately affect students' academic performance, particularly in demanding subjects such as physics (Ferry D, 2025).

Numerous studies have reported that physics examinations frequently become a source of psychological pressure for students because they are required not only to understand complex scientific concepts but also to apply those concepts accurately to solve problems under time constraints (Lestari et al., 2015; Rahmawati & Sari, 2025). One of the most common psychological responses arising in such situations is test anxiety (Jarso et al., 2023). Test anxiety is more than a temporary feeling of nervousness before an examination; rather, it is a multidimensional psychological construct encompassing cognitive, emotional, physiological, and behavioral components that may interfere with concentration, memory retrieval, cognitive processing, and decision-making during examinations (Nazir et al., 2021; Prasetyaningtyas et al., 2022; Revilla, 2009). Within the context of physics education, this phenomenon is commonly referred to as physics test anxiety, which specifically describes the anxiety experienced by students when they are evaluated through physics examinations (Molin et al., 2021).

The relationship between test anxiety and academic achievement has attracted considerable research attention over the past several decades. Previous studies have consistently demonstrated that higher levels of test anxiety are negatively associated with students' academic performance because anxiety reduces working memory capacity, impairs concentration, and diminishes the effectiveness of problem-solving processes. A comprehensive meta-analysis conducted by von der Embse et al. (2018) confirmed that test anxiety is significantly and negatively associated with academic achievement across various educational levels. These findings are further supported by Fadli et al. (2024), who reported a significant negative relationship between anxiety and physics learning achievement among tenth-grade students in Indonesia. Similarly, Kusyairy, (2014) found that students experiencing higher levels of anxiety tended to encounter greater difficulties in learning physics and consequently achieved lower academic performance. Collectively, these findings suggest that test anxiety should not merely be regarded as an individual psychological issue but rather as an important educational factor that substantially influences the quality of physics learning and students' ability to demonstrate their actual competence during physics assessments.

The emergence of academic anxiety can be theoretically explained through the Control-Value Theory of Achievement Emotions proposed by Pekrun (2006). According to this theory, anxiety arises when students perceive an academic task as highly important *(high value)* while simultaneously believing that they have limited control over successfully completing the task (*low perceived control*). Consequently, the learning environment plays a crucial role in shaping students' perceptions of both the value of academic tasks and their perceived competence in dealing with examinations. Differences in school characteristics, academic culture,

instructional practices, teacher support, and the values promoted within schools may therefore lead to different emotional experiences when students encounter academic assessments (Embse & Hasson, 2012). From this perspective, school context represents an important theoretical factor that may explain variations in physics test anxiety among students.

Within the Indonesian education system, senior secondary education is generally divided into two major types of institutions, namely public senior high schools and religious-based senior high schools. Religious-based senior high schools integrate the national curriculum with religious education, providing students with learning experiences that differ from those of their counterparts in public schools (Musfah et al., 2021). In contrast, public senior high schools primarily implement the national curriculum, which emphasizes general academic subjects and competency development in accordance with government educational policies (Mulyana, 2025). These differences in curriculum structure and educational environment may influence how students perceive academic demands, prepare for examinations, and respond to learning assessments. Despite these contextual differences, research on physics test anxiety in Indonesia has predominantly focused on examining the relationship between anxiety and academic achievement or identifying individual factors associated with anxiety, while little attention has been paid to whether students from different school contexts experience physics test anxiety differently. For example, Putranta & Jumadi (2019) investigated physics test anxiety among students in a single religious-based senior high school, but their findings did not allow comparisons with students from public senior high schools. Consequently, empirical evidence regarding the influence of school context on physics test anxiety among Indonesian secondary school students remains limited.

In addition to these contextual limitations, previous studies also exhibit important methodological limitations. Most investigations of test anxiety have employed the Classical Test Theory (CTT) approach, in which the psychometric properties of an instrument are inherently dependent on the characteristics of the sample from which the data are obtained (Hambleton & Jones, 2005). Such dependency poses a challenge when comparing different groups because it does not adequately establish whether individual questionnaire items function equivalently across populations. In contrast, Rasch Measurement Theory provides sample-independent parameter estimation while enabling the application of Differential Item **Functioning (DIF)** analysis to evaluate whether measurement items operate equivalently across groups (Bond & Fox, 2015). Therefore, DIF analysis serves not only as a method for comparing students' levels of physics test anxiety but also as a psychometric procedure for ensuring that observed differences between groups reflect genuine differences in the latent construct rather than measurement bias (Pallant & Tennant, 2007).

Despite the growing body of research on test anxiety, important gaps remain in understanding whether physics test anxiety differs across educational contexts and whether the measurement instrument functions equivalently when applied to different student populations. To address these gaps, the present study compares physics test anxiety between students enrolled in religious-based and public senior high schools in Indonesia using the **Test** Anxiety Inventory for Physics (TAIP) developed by Cioffi et al. (2024), which was adapted for Indonesian secondary school students. Furthermore, rather than relying solely on comparisons of anxiety scores, this study evaluates the measurement equivalence of the adapted instrument through Differential Item Functioning (DIF) analysis within the Rasch Measurement framework. Establishing measurement equivalence before comparing latent constructs is essential because observed differences between groups may otherwise reflect measurement bias rather than genuine differences in the latent construct (Zumbo, 1999). Consequently, this study

contributes not only by providing empirical evidence regarding variations in physics test anxiety across different school contexts but also by strengthening the psychometric validity of cross-group comparisons using a modern measurement approach.

Accordingly, this study aims to compare the levels of physics test anxiety between students attending religious-based and public senior high schools and to identify potential differences in item functioning across school contexts using Differential Item Functioning (DIF) analysis within the Rasch Measurement framework. The findings are expected to provide a more comprehensive understanding of how school context relates to students' physics test anxiety while also providing psychometric evidence regarding the measurement equivalence of the adapted Test Anxiety Inventory for Physics (TAIP). Ultimately, the findings may assist educators in designing learning and assessment strategies that are more responsive to students' psychological conditions and support the use of valid and equitable instruments for measuring physics test anxiety in future educational research.

2. Method

This study employed a cross-sectional survey using a quantitative approach to compare the levels of physics test anxiety between students from public senior high schools and religious-based senior high schools and to identify potential differences in item functioning across school contexts using Differential Item Functioning (DIF) analysis within the Rasch Measurement Theory framework (Kurniawati & Rindrayani, 2025). A survey approach was considered appropriate because it enables efficient collection of students' perceptions through a structured questionnaire and facilitates comparisons between different educational contexts.

The study involved 236 senior high school students from several public senior high schools and religious-based senior high schools located in the Special Region of Yogyakarta and Central Java, Indonesia. The participants consisted of students from Grades 10, 11, and 12, including 94 students from public senior high schools and 142 students from religious-based senior high schools. Based on gender, the respondents comprised 116 male and 120 female students.

Participants were selected using convenience sampling, a sampling technique based on the accessibility of respondents who were willing to participate in the study (Golzar & Noor, 2022). This technique was chosen because data collection depended on the willingness of schools to grant permission for the study as well as the voluntary participation of teachers and students in completing the online questionnaire. Convenience sampling is widely used in educational survey research employing online questionnaires (Regmi et al., 2016). Consequently, the number of participants in each group was not predetermined but depended on the number of eligible students who voluntarily participated during the data collection period. Therefore, the numbers of respondents from public and religious-based senior high schools were not equal. Accordingly, the findings of this study should be interpreted with caution, as their generalizability is limited to populations with characteristics similar to those of the participants included in this study.

Following data collection, data screening was conducted using the Rasch model. Four respondents were identified as having extreme person measures, resulting in unstable logit estimates that could not be reliably estimated by the Rasch model. Consequently, these four respondents were excluded from further analysis, leaving 232 respondents for the final analysis.

The instrument used in this study was the Test Anxiety Inventory for Physics (TAIP) developed by Cioffi et al. (2024).The TAIP was selected because it was specifically designed

to measure physics test anxiety, making it more appropriate than general anxiety instruments that do not specifically assess anxiety related to physics examinations.

For the purposes of this study, the instrument was translated from English into Indonesian using forward translation and back translation procedures to maintain semantic equivalence between the original and translated versions. First, all questionnaire items were translated into Indonesian (*forward translation*). Subsequently, the translated version was translated back into English (*back translation*) and compared with the original instrument to ensure that the meaning of each item was retained before the questionnaire was administered to participants. The TAIP consists of 20 items distributed across four dimensions: Worry (WR), Emotionality (EM), Interference (IF), and Lack of Confidence (LC). Each dimension comprises five items representing different aspects of students' anxiety toward physics examinations. Details of each dimension and the corresponding number of items are presented in Table 1.

Table 1. Distribution of the TAIP Instrument

No	Original version	Adapted version
	Worry subscale (WR)	
1	I think about how important the examination is for me.	*Saya berpikir tentang betapa pentingnya tes ini bagi saya.*
2	I think about how important it is for me to receive a good result.	*Saya berpikir tentang betapa pentingnya bagi saya untuk mendapatkan hasil yang baik.*
3	I worry about my results.	*Saya khawatir tentang hasil yang akan saya peroleh.*
4	I am concerned about my grades.	*Saya merasa cemas tentang nilai saya.*
5	I think about what will happen if I don't do well.	*Saya memikirkan apa yang akan terjadi jika saya tidak mendapatkan hasil yang baik.*
	Emotionality subscale (ES)	
1	I feel my heart beating fast.	*Saya merasa jantung saya berdebar kencang.*
2	I feel anxious.	*Saya merasa cemas.*
3	I tremble with fear.	*Saya merasa gemetar karena takut.*
4	I feel somewhat overwhelmed.	*Saya merasa sedikit kewalahan.*
5	I feel upset.	*Saya merasa kesal.*
	Interference subscale (IS)	
1	Distracting thoughts keep "popping" into my head.	*Pikiran-pikiran yang mengganggu terus muncul di kepala saya.*
2	I am preoccupied by other thoughts that distract me.	*Saya terganggu oleh pikiran-pikiran lain yang mengalihkan perhatian saya.*
3	I easily lose my train of thoughts.	*Saya mudah kehilangan fokus dalam pikiran saya.*
4	I forget things because I am too preoccupied with my personal problems.	*Saya sering lupa banyak hal karena terlalu sibuk dengan masalah pribadi saya.*
5	My concentration is interrupted by interfering thoughts.	*Konsentrasi saya terganggu oleh pikiran yang mengganggu.*
	Lack of Confidence subscale (LCS)	
1	I am confident about my performance.	*Saya percaya diri dengan kinerja saya.*
2	I have faith in my own performance.	*Saya yakin dengan kemampuan saya sendiri.*
3	I am satisfied with myself.	*Saya merasa puas dengan diri saya sendiri.*
4	I think that I will succeed.	*Saya berpikir bahwa saya akan berhasil.*
5	I am convinced that I will do well.	*Saya yakin bahwa saya akan sukses.*

Aspect “The Worry dimension reflects students' cognitive concerns regarding possible failure or unsatisfactory examination results. Emotionality represents emotional and physiological reactions, such as tension, nervousness, and anxiety experienced during examinations. The Interference dimension indicates the extent to which anxiety disrupts students' concentration and thinking processes while completing the examination, including difficulty concentrating, intrusive thoughts, and difficulty recalling previously learned information. Meanwhile, Lack of Confidence reflects students' confidence in their ability to understand physics concepts and solve examination problems.

Responses were recorded using a five-point Likert scale, consisting of Never, Rarely, Sometimes, Often, and Always (see Table 2). This scale allows respondents to indicate the frequency with which they experience each anxiety-related situation, thereby enabling systematic measurement of physics test anxiety.

Table 2. Questionnaire Response Conversion

Answer Choices	Answer Score
Almost Never	1
Rarely	2
Often	3
Almost Always	4
Everytime	5

Data collection began after obtaining permission from the participating schools. Subsequently, the researchers coordinated with physics teachers to determine an appropriate schedule for questionnaire administration.

The questionnaire was distributed online using Google Forms. Before completing the questionnaire, all participants received information regarding the purpose of the study and were informed that participation was entirely voluntary. Students were also informed that there would be no academic consequences if they decided not to participate. In addition, respondents' identities were not collected during the analysis process, ensuring the confidentiality of the collected data. These procedures were implemented to encourage honest responses and minimize potential response bias. The average time required to complete the questionnaire ranged from 10 to 15 minutes.

Data analysis was performed using Winsteps for Rasch analysis and IBM SPSS Statistics version 23 for statistical analysis. Rasch analysis was employed to estimate respondents' levels of physics test anxiety (person measures) and to ensure that the collected data met the requirements for subsequent Rasch-based analyses before estimating person measures and conducting Differential Item Functioning (DIF) analysis. The Rasch approach was selected because it provides parameter estimates that are relatively independent of sample characteristics and enables the evaluation of item functioning across different respondent groups (Bond & Fox, 2015)

As part of the Rasch analysis, item fit was examined using Infit Mean Square (Infit MNSQ) and Outfit Mean Square (Outfit MNSQ) statistics. Items were considered to fit the Rasch model if their MNSQ values ranged between 0.50 and 1.50 (Bond & Fox, 2015). Subsequently, the consistency of the measurement model was examined using Cronbach's alpha, person reliability, and item reliability. The interpretation criteria for these indicators are presented in Table 3.

Table 3. Reliability and Validity Criteria in Rasch Analysis

Statistic	Index	Interpretation
Cronbach's Alpha (KR-20)	<0.5	Low
	0.5 - <0.6	Moderate
	0.6 - < 0.7	Good
	0.7 – 0.8	High
	> 0.8	Very High
Item and Person Reliability	> 0.67	Low
	0.67 - 0.80	Fair
	0.81 - 0.90	Good
	0.91- 0.94	Very Good
	0.94	Excellent
Outfit & Infit MNSQ	0.50 – 1.50	Fit

(Source: Faradillah et al. (2021 as cited in Boone et al., 2014))

The Rasch model was then used to estimate the Logit Value of Person (LVP) for each respondent, representing individual levels of physics test anxiety. These logit values were visualized using a Wright Map and classified into four categories: very high, high, low, and very low.

The Logit Value of Person (LVP) scores were subsequently analysed using the Mann–Whitney U test to examine differences in physics test anxiety between students from public senior high schools and religious-based senior high schools. Differences were considered statistically significant when the p-value was less than 0.05.

Finally, Differential Item Functioning (DIF) analysis was conducted to determine whether individual TAIP items functioned differently across the two school contexts. DIF analysis was performed for all four dimensions of the instrument, namely Worry (WR), Emotionality (EM), Interference (IF), and Lack of Confidence (LC). An item was considered to exhibit Differential Item Functioning when its p-value was less than 0.05 (Handayani et al., 2023). This analysis was conducted to ensure that any observed differences in physics test anxiety between the two groups reflected genuine differences in the measured construct rather than differences caused by item functioning across school contexts

3. Result and Discussion

This section presents the results of the analyses conducted to examine students' physics test anxiety using the Test Anxiety Inventory for Physics (TAIP). As part of the Rasch analysis, item fit and measurement consistency were first examined before estimating person measures. The resulting person measures were subsequently used to classify students' levels of physics test anxiety, compare anxiety levels between public senior high schools and religious-based senior high schools using the Mann–Whitney U test, and evaluate item functioning across school contexts through Differential Item Functioning (DIF) analysis.

a. Item Fit Analysis of the TAIP Instrument

The results of the Rasch item fit analysis are presented in Table 4. Item fit was evaluated using Infit Mean Square (Infit MNSQ) and Outfit Mean Square (Outfit MNSQ) statistics.

Table 4. Results of Rasch Item Fit for the TAIP Instrument

No	Item Code	Infit MNSQ	Outfit MNSQ	Status (Fit/Misfit)
1	WR1	1.29	1.32	Fit
2	WR2	1.02	1.24	Fit
3	WR3	0.89	0.92	Fit
4	WR4	0.84	0.87	Fit
5	WR5	0.77	0.80	Fit
6	ES1	0.90	0.90	Fit
7	ES2	0.85	0.84	Fit
8	ES3	1.21	1.14	Fit
9	ES4	0.76	0.75	Fit
10	ES5	0.93	0.91	Fit
11	IS1	0.91	0.89	Fit
12	IS2	0.78	0.76	Fit
13	IS3	0.71	0.72	Fit
14	IS4	0.97	1.02	Fit
15	IS5	0.76	0.78	Fit
16	LCS1	1.16	1.24	Fit
17	LCS2	1.38	1.50	Fit
18	LCS3	1.48	1.55	Fit
19	LCS4	1.27	1.38	Fit
20	LCS5	1.20	1.29	Fit

As shown in Table 4, nineteen of the twenty items exhibited Infit MNSQ and Outfit MNSQ values within the recommended range of 0.50–1.50, indicating satisfactory fit to the Rasch model. One item (LCS3) showed an Outfit MNSQ value of 1.55, which was only slightly above the recommended upper limit, whereas its Infit MNSQ value remained within the acceptable range (1.48).

Although the Outfit MNSQ value of LCS3 slightly exceeded the recommended criterion, the item was retained for subsequent analyses for several reasons. First, the deviation from the recommended upper limit was minimal (0.05 above the criterion). Second, the Infit MNSQ value remained within the acceptable range, indicating that the response pattern for this item was still consistent with the expectations of the Rasch model. In Rasch analysis, Infit statistics are generally considered more informative than Outfit statistics because they are more sensitive to response patterns on items targeted to respondents' estimated ability levels, whereas Outfit statistics are more influenced by unexpected responses on items that are relatively easy or difficult (Bond & Fox, 2015; Linacre, 2002). Therefore, the slightly elevated Outfit value observed for LCS3 was not considered sufficient to justify removing the item. Considering the overall pattern of item fit, all twenty TAIP items were retained for subsequent analyses. The results indicate that the instrument functioned adequately for measuring physics test anxiety in the present sample and was suitable for estimating person measures and conducting Differential Item Functioning (DIF) analysis.

b. Reliability Analysis

The reliability analysis showed that the TAIP instrument demonstrated satisfactory measurement consistency. As presented in **Table 3**, the Cronbach's Alpha (KR-20) coefficient was 0.86, indicating a high level of internal consistency among the questionnaire items. According to Boone et al. (2014), Cronbach's alpha values above 0.80 indicate very good interaction between respondents and items.

The person reliability value was 0.83, suggesting that the instrument was able to consistently distinguish respondents according to their levels of physics test anxiety. Meanwhile, the item reliability value reached 0.99, indicating excellent stability of item difficulty estimates across the respondent sample (Boone et al., 2014). These findings suggest that the hierarchy of item difficulty was highly stable and would likely be replicated if administered to respondents with similar characteristics.

Overall, the reliability indices indicate that the TAIP instrument provided sufficiently consistent measurements for estimating person measures and conducting subsequent comparative analyses.

c. Distribution of Physics Test Anxiety Based on Logit Value of Person (LVP)

The distribution of students' physics test anxiety was estimated using the Logit Value of Person (LVP) generated from the Rasch model. The LVP represents each respondent's level of physics test anxiety on a logit scale, where higher logit values indicate higher levels of anxiety toward physics examinations.

Figure 1 presents the distribution of respondents based on their LVP values. To facilitate interpretation, respondents were classified into four anxiety categories according to the mean person measure and one standard deviation from the mean. Based on the mean logit value of **0.04** and a standard deviation of **0.60**, students' physics test anxiety was categorized as **very high** ($X \geq 0.64$), **high** ($0.04 \leq X < 0.64$), **low** ($-0.56 \leq X < 0.04$), and **very low** ($X < -0.56$).

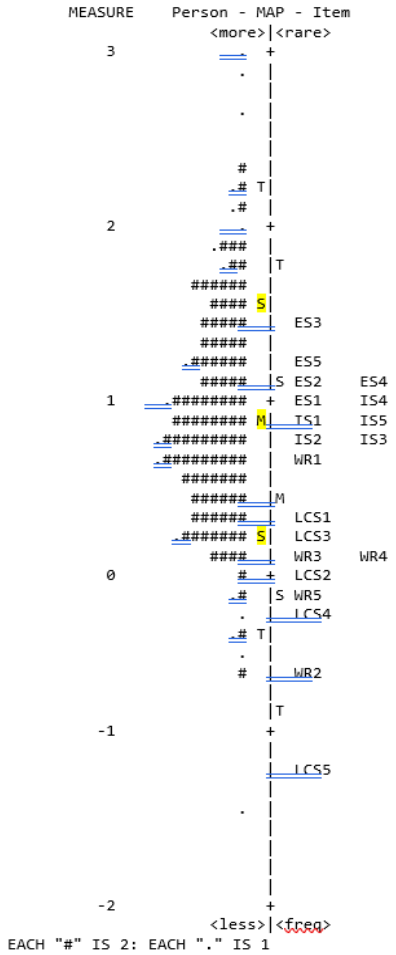

Figure 1. Distribution of students' physics test anxiety based on the Logit Value of Person (LVP)

The classification results indicate that most students were categorized as having low levels of physics test anxiety, accounting for 88 students (37.93%). This was followed by the high category with 63 students (27.16%), the very low category with 53 students (22.84%), and the very high category with 28 students (12.07%). These findings suggest that, overall, the majority of respondents experienced relatively low to moderate levels of anxiety when taking physics examinations.

To further illustrate the distribution of students' physics test anxiety according to school type, Wright Maps were generated separately for religious-based senior high schools and public senior high schools, as presented in **Figure 2** and **Figure 3**, respectively. These figures provide a descriptive comparison of the distribution of person measures across the two school contexts prior to the statistical comparison using the Mann–Whitney U test.

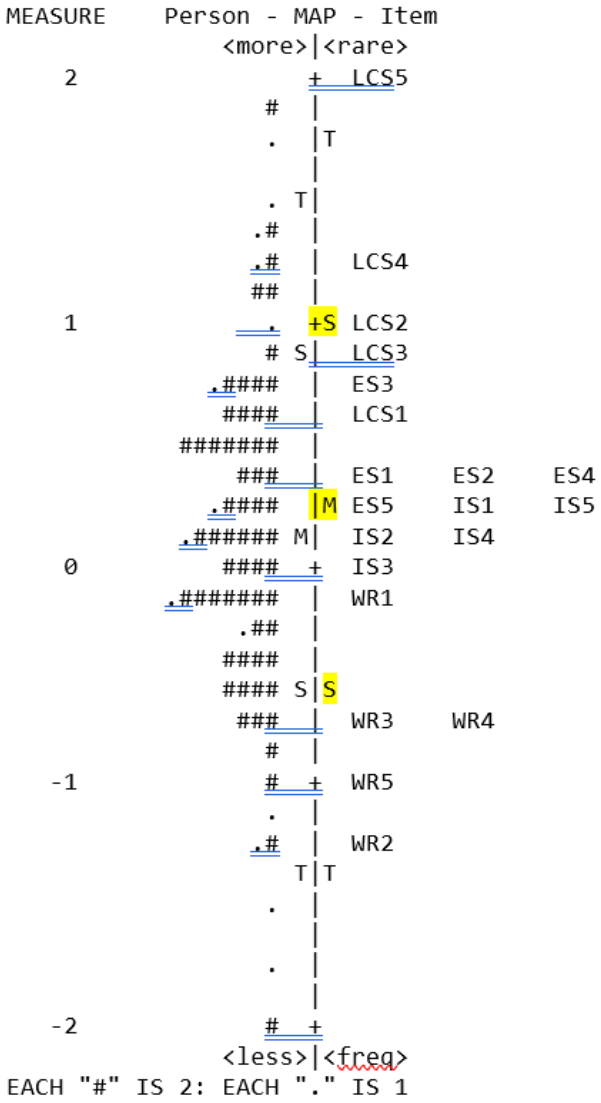

Figure 2. Wright Map of Students' Anxiety Levels on Physics Tests at Religion-Based High Schools

Using the same classification procedure described previously, students from **religious-based** senior high schools were categorized according to the subgroup-specific mean person measure and one standard deviation. Based on a mean logit value of 0.06 and a standard deviation of 0.66, students' physics test anxiety was classified into four categories: very high

(X ≥ 0.72), high (0.06 ≤ X < 0.72), low (−0.60 ≤ X < 0.06), and very low (X < −0.60). The distribution of students' physics test anxiety is presented in Figure 2. The largest proportion of students was classified into the low anxiety category (50 students; 36.23%), followed by the high anxiety category (39 students; 28.26%), the very low anxiety category (34 students; 24.64%), and the very high anxiety category (15 students; 10.87%). These descriptive results provide an overview of the distribution of physics test anxiety among students in religious-based senior high schools prior to comparison with students from public senior high schools.

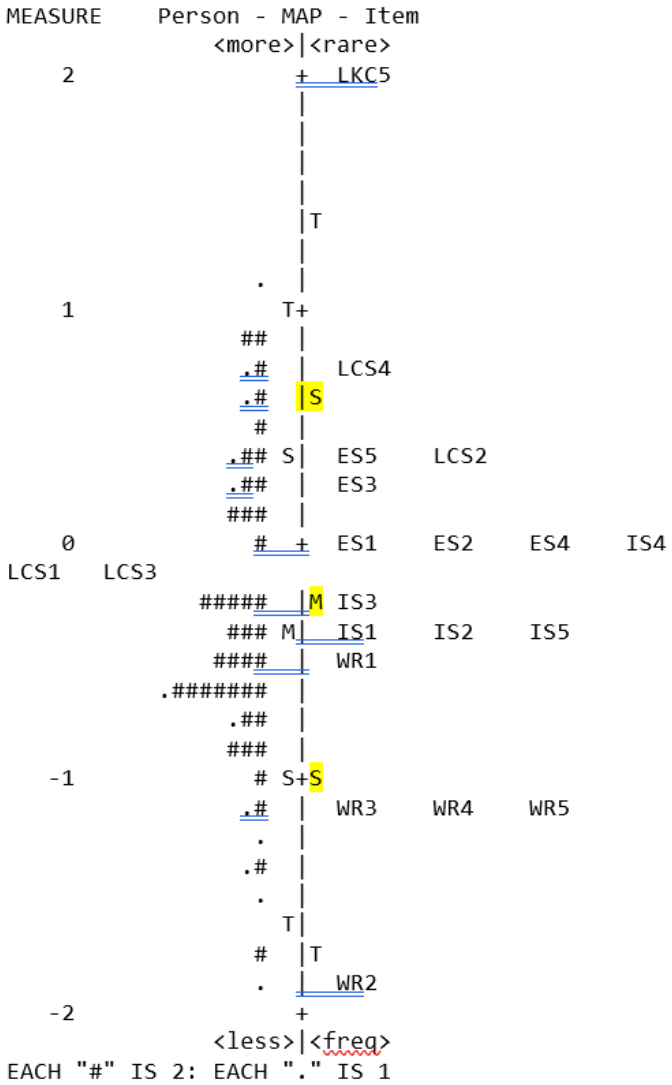

Figure 3. Wright Map of Anxiety Levels among General Senior High School Students

Anxiety levels among students at public high schools regarding physics exams were also grouped into four categories: very high (X ≥ 0.63), high (0.06 ≤ X < 0.63), low (−0.51 ≤ X < 0.06), and very low (X < −0.51), based on a mean of 0.06 and a standard deviation of 0.57. Analysis of individual measurements showed that the majority of students fell into the low anxiety category, namely 43 students (45.74%), followed by the very low category with 20 students (21.28%), the high category with 18 students (19.15%), and the very high category with 13 students (13.83%).

A descriptive comparison between the two school groups indicates that students from both public and religious-based senior high schools were predominantly classified into the low

anxiety category. However, the combined proportion of students in the low and very low anxiety categories was higher in public senior high schools (67.02%) than in religious-based senior high schools (60.87%). In contrast, religious-based senior high schools showed a higher proportion of students in the high anxiety category (28.26%) than public senior high schools (19.15%). The proportion of students classified into the very high anxiety category was relatively similar in both groups, accounting for 10.87% in religious-based senior high schools and 13.83% in public senior high schools. These descriptive findings indicate a tendency toward differences in the distribution of physics test anxiety between students from public and religious-based senior high schools. However, descriptive comparisons alone cannot determine whether these differences are statistically significant. Therefore, a Mann–Whitney U test was subsequently performed to examine whether the observed differences in students' physics test anxiety between the two school contexts were statistically significant.

d. Mann-Whitney test

The Logit Value of Person (LVP) obtained from Rasch modeling was used as an indicator of students' anxiety levels regarding the physics test. Next, to test whether there were significant differences in anxiety levels between students at public high schools and religious high schools, a Mann–Whitney test was conducted using the LVP logit values. The test results are presented in **Table 5.**

Table 5. Mann-Whitney Test Results

	Total
Mann–Whitney U	6096,500
Z	−0,776
Asymp. Sig. (2-tailed)	0,437

Results from the Mann–Whitney test indicate that there is no significant difference in the level of anxiety related to physics exams between students at public high schools and religious high schools ($p = 0.437 > 0.05$). These results indicate that the observed difference in anxiety levels between the two types of schools is not statistically significant. Thus, school type is not a major factor distinguishing students' anxiety levels regarding physics exams.

One possible reason for this is the use of similar learning and assessment systems in both types of schools, resulting in relatively comparable levels of academic pressure experienced by students. Furthermore, test anxiety is more strongly influenced by various internal factors, such as self-confidence, prior test experience, learning motivation, study discipline, and science literacy (Asmawati et al., 2021; Purnama et al., 2023; Rahmani et al., 2024). Therefore, the differences in characteristics between public high schools and religious high schools are not yet strong enough to produce statistically significant differences in the level of physics test anxiety. Nevertheless, a more in-depth comparison can be conducted through Rasch modeling using Differential Item Functioning (DIF) analysis to identify possible differences in responses to physics test anxiety items in both types of schools.

e. Differential Item Functioning (DIF) Test for Religion-Based High Schools and Public High Schools

The results of the Differential Item Functioning (DIF) analysis between students from religion-based high schools and public high schools are presented in **Figure 4.**

DIF class/group specification is: DIF=$S4W1

Person CLASSES	SUMMARY DIF CHI-SQUARED	D.F.	PROB.	BETWEEN-CLASS/GROUP Item UNWTD MNSQ	ZSTD	Number	Name
2	.0000	1	1.0000	.0291	-1.00	1	WR1
2	6.8135	1	.0090	7.0068	2.41	2	WR2
2	.0000	1	1.0000	.0010	-1.44	3	WR3
2	.4329	1	.5106	.4337	-.04	4	WR4
2	.3709	1	.5425	.3705	-.13	5	WR5
2	.6704	1	.4129	.6717	.21	6	ES1
2	.5431	1	.4612	.5432	.08	7	ES2
2	.7579	1	.3840	.7622	.29	8	ES3
2	.0000	1	1.0000	.0047	-1.29	9	ES4
2	10.0563	1	.0015	10.3825	2.98	10	ES5
2	.7185	1	.3966	.7196	.25	11	IS1
2	.4237	1	.5151	.4223	-.06	12	IS2
2	.4863	1	.4856	.4870	.02	13	IS3
2	.1942	1	.6594	.1930	-.42	14	IS4
2	1.3726	1	.2414	1.3735	.71	15	IS5
2	1.6472	1	.1993	1.6479	.86	16	LCS1
2	.2876	1	.5918	.2864	-.25	17	LCS2
2	8.1832	1	.0042	8.4050	2.66	18	LCS3
2	.2386	1	.6252	.2387	-.33	19	LCS4
2	3.7313	1	.0534	3.8007	1.66	20	LCS5

TABLE 30.5 Olah data TAIP_03_05_26 ZOU219WS.TXT May 3 2026 20:36
INPUT: 232 Person 20 Item REPORTED: 232 Person 20 Item 5 CATS WINSTEPS 4.6.1

Figure 4. DIF Analysis between Religion-Based High Schools and Public High Schools

Based on **Figure 4**, showed that 17 of the 20 TAIP items functioned equivalently across the two school groups, indicating that most questionnaire items measured physics test anxiety consistently regardless of school type. However, three items showed DIF. These were WR2 ***"Saya berpikir tentang betapa pentingnya bagi saya untuk mendapatkan hasil yang baik"***, ES5 ***"Saya merasa kesal"***, dan LCS3 ***"Saya merasa puas dengan diri saya sendiri".*** These findings indicate that although students from the two school contexts exhibited comparable overall levels of physics test anxiety, several specific questionnaire items were interpreted or responded to differently.

The identification of only three DIF items suggests that the TAIP instrument generally functioned consistently across both school contexts, supporting its use for comparing physics test anxiety between students from religious-based and public senior high schools. At the same time, the presence of DIF in several items indicates that similarities in overall anxiety levels do not necessarily imply identical response patterns for every dimension of physics test anxiety. This finding complements the previous Mann–Whitney U test, which showed no significant difference in overall anxiety levels, by demonstrating that differences between school contexts may emerge at the level of specific anxiety indicators rather than at the overall construct level.

To further examine these differences, response patterns for the three DIF items are presented in **Figure 5.**

The response pattern for WR2 showed that students from public senior high schools had lower logit values than students from religious-based senior high schools, indicating that public school students were more likely to endorse the statement at the same overall anxiety level. This finding suggests that students from the two school contexts may assign different levels of importance to obtaining good academic results when responding to this item. According to the

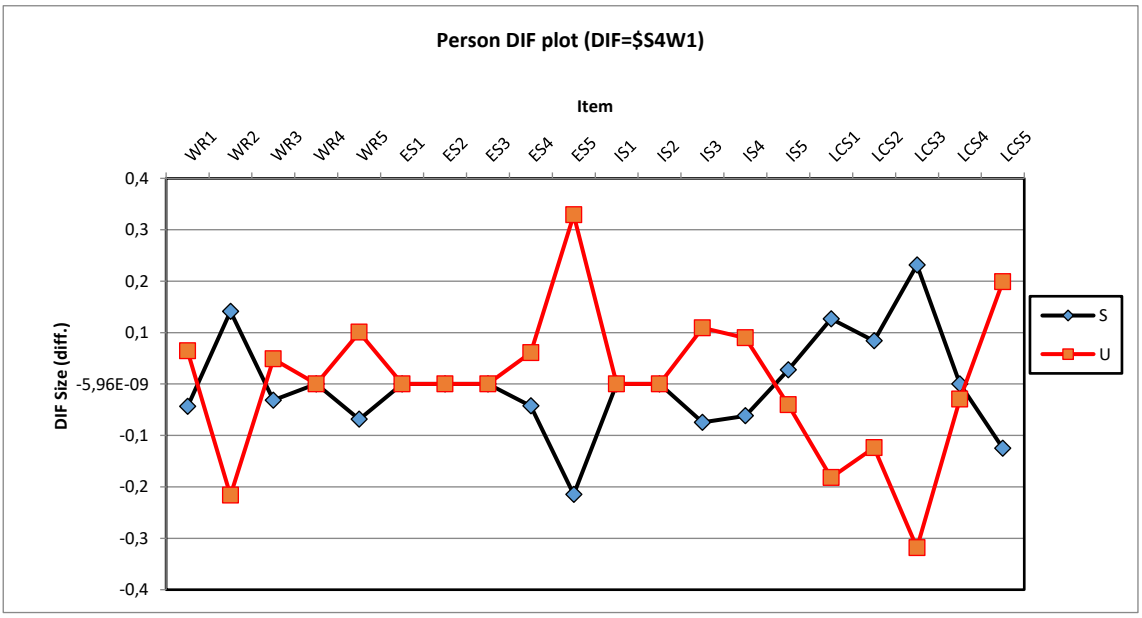

Figure 5. Response Patterns of DIF for Public High Schools and Religion-Based High Schools

Control–Value Theory of Achievement Emotions (Pekrun, 2006), tudents' emotional responses toward examinations are influenced by the value they attach to academic success and their perceived control over achieving it. Therefore, differences in responses to WR2 may reflect differences in how students evaluate the importance of academic achievement rather than differences in their overall levels of physics test anxiety. Similar findings have been reported by Ifdil & Ardi (2013), who suggested that students' perceptions of academic achievement are associated with academic anxiety. However, because the present study did not directly measure students' academic orientation or achievement motivation, this interpretation should be considered tentative.

For ES5, students from religious-based senior high schools exhibited lower logit values than students from public senior high schools, indicating that they were more likely to agree with the statement "I feel upset" despite having comparable overall anxiety levels. This result suggests that students from the two school contexts may express the emotional dimension of physics test anxiety differently. Rather than indicating higher overall anxiety, the finding reflects differences in responses to a specific emotional indicator. Previous studies have also reported that educational environments may influence students' emotional responses to academic evaluation (Iskandar et al., 2018) Nevertheless, because the present study did not directly examine emotional regulation, school culture, or other contextual variables, the underlying reasons for this difference cannot be determined from the available data.

For LCS3, students from public senior high schools showed lower logit values than students from religious-based senior high schools, indicating that they were more likely to agree with the statement "*Saya merasa puas dengan diri saya sendiri*" at the same overall anxiety level. This finding suggests that the two groups may differ in how they interpret self-confidence within the context of physics examinations. However, the present study did not directly assess students' self-confidence or related psychological constructs. Therefore, although previous studies have suggested that educational environments may contribute to students' self-evaluation (Januaripin & Munasir, 2024; Muhaimin, 2023), the influence of these factors cannot be confirmed in the present study and should be interpreted cautiously.

Overall, the findings demonstrate that although students from religious-based and public senior high schools exhibited comparable overall levels of physics test anxiety, they responded differently to several specific aspects of the TAIP instrument. This result highlights the importance of conducting Differential Item Functioning (DIF) analysis when comparing psychological constructs across different educational contexts, as equivalent total scores do not necessarily indicate equivalent interpretations of individual questionnaire items. Consequently, DIF analysis provides additional evidence that complements comparisons based solely on total anxiety scores and contributes to a more accurate interpretation of students' physics test anxiety across different school settings.

4. Conclusion

This study compared physics test anxiety between students from religious-based senior high schools and public senior high schools and examined potential differences in item functioning using Differential Item Functioning (DIF) analysis within the Rasch measurement framework. The findings showed that students from both school types generally exhibited low levels of physics test anxiety, and the Mann–Whitney U test indicated no statistically significant difference in overall anxiety levels between the two groups. However, the DIF analysis identified three items (WR2, ES5, and LCS3) that functioned differently across school contexts, indicating that although students demonstrated comparable overall levels of physics test anxiety, they responded differently to several specific aspects of the instrument.

The scientific contribution of this study lies in demonstrating that comparisons based solely on overall anxiety scores may overlook meaningful differences in students' responses to individual questionnaire items. By combining group comparison analysis with Differential Item Functioning (DIF) analysis, this study provides empirical evidence that school context may influence responses to particular dimensions of physics test anxiety, even when overall anxiety levels are statistically comparable. These findings highlight the importance of evaluating item functioning in comparative educational research to ensure that conclusions are based not only on total scores but also on the equivalence of measurement across groups.

From a practical perspective, the findings suggest that educators and researchers should consider potential differences in students' responses to specific aspects of physics test anxiety when designing assessment strategies, interpreting questionnaire results, and providing academic support. Furthermore, the results support the use of DIF analysis as a complementary approach for evaluating the fairness and comparability of psychological measurements across different educational contexts. This study has several limitations. The Indonesian version of the Test Anxiety Inventory for Physics (TAIP) was adapted using forward and backward translation procedures without conducting a comprehensive cultural adaptation involving expert review or curriculum-based content evaluation. In addition, this study did not directly measure contextual variables such as school culture, religiosity, academic orientation, or social support that may contribute to differences in students' responses to specific anxiety items. Future research is therefore recommended to conduct a more comprehensive cross-cultural adaptation of the TAIP instrument and to incorporate contextual and psychological variables that may explain the observed DIF patterns across different educational settings.

5. References

Asmawati, A. A., Sugeng, S., & Labulan, P. M. (2021). Pengaruh Disiplin Belajar,Kecemasan dan Perhatian Orang Tua Terhadap Hasil belajar Matematika Siswa. *Jurnal Primatika*,

10(1), 1–10.

Bobur, S., Otabek, N., & Guzal, A. (2021). The role of physics in the study of nature and its role in development with other sciences. *International Journal of Engineering in Computer Science*, *3*(2), 5–8. https://doi.org/10.33545/26633582.2021.v3.i2a.50

Bond, T., & Fox, C. (2015). *Applying the Rasch Model; Fundamental Measurement in the Human Sciences*.

Boone, W., Staver, J., & Yale, M. (2014). *Rasch Analysis in the Human Sciences* (1st ed.). Springer Dordrecht. https://doi.org/10.1007/978-94-007-6857-4

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

Embse, N. von der, & Hasson, R. (2012). Test Anxiety and High-Stakes Test Performance Between School Settings: Implications for Educators. *Preventing School Failure: Alternative Education for Children and Youth*, *56*(3), 180–187. https://doi.org/10.1080/1045988X.2011.633285

Fadli, O. M., Rusdy, R., & Wahid, M. A. (2024). Pengaruh Tingkat Kecemasan terhadap Hasil Belajar Siswa Kelas X pada Pelajaran Fisika di SMA Inshafuddin Banda Aceh. *Jurnal Ilmiah Mahasiswa*, *2*, 52–66.

Faradillah, A., Jusra, H., & Pradipta, T. R. (2021). Validity and Reliability Analysis of Algebraic Reasoning Test Instrument. *Journal of Physics: Conference Series PAPER*. https://doi.org/10.1088/1742-6596/1933/1/012116

Ferry D, W. (2025). Eksplorasi Tingkat Kecemasan Artificial Intelligence (AI) pada Mahasiswa Pendidikan Biologi di Institut Agama Islam Negeri Kerinci. *Edusainstika: Jurnal Pembelajaran MIPA*, *5*(2), 45–59.

Golzar, J., & Noor, S. (2022). Defining Convenience Sampling in a Scientific Research. *International Journal of Education and Language Studies*, *1*(December), 72–77.

Hambleton, R., & Jones, R. (2005). An NCME Instructional Module on Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*, *12*, 38–47. https://doi.org/10.1111/j.1745-3992.1993.tb00543.x

Handayani, T., Sukarelawan, M. I., Effendi, M., & Mohd, E. (2023). Adaptation and evaluation of psychometric properties of HELAM using the Rasch model in the Indonesian context. *International Journal of Evaluation and Research in Education*, *12*(4), 2362–2373. https://doi.org/10.11591/ijere.v12i4.27703

Ifdil, T., & Ardi, Z. (2013). Kondisi Stres Akademik Siswa SMA Negeri di Kota Padang. *Jurnal Konseling Dan Pendidikan*.

Iskandar, B. J., Noupal, M., & Setiawan, K. C. (2018). Sikap Tawakal dengan kecemasan Menghadapi Ujian Nasional Pada Siswa Kelas XII Madrasah Aliyah di Kota Palembang. *Jurnal Psikologi Islami*, *4*(1). https://doi.org/10.19109/psikis.v4i1.2158

Januaripin, M., & Munasir, M. (2024). Kepercayaan Diri sebagai Predikator Prestasi Akademik Siswa. *Kamaliyah : Jurnal Pendidikan Agama Islam*, *2*(1), 114–128.

https://doi.org/10.69698/jpai.v2i1.575

Jarso, M. H., Tariku, M., Mamo, A., Tsegaye, T., & Gezimu, W. (2023). Test anxiety and its determinants among health sciences students at Mattu University: a cross-sectional study. *Frontiers in Psychiatry*, *14*. https://doi.org/10.3389/fpsyt.2023.1241940

Kurniawati, E., & Rindrayani, S. R. (2025). Pendekatan Kuantitatif dengan Penelitian Survei : Studi Kasus dan Implikasinya. *Sosial : Jurnal Ilmiah Pendidikan IPS*, *3*(1). https://doi.org/10.62383/sosial.v3i1.596

Kusyairy, U. (2014). Hubungan Kecemasan dan Kesulitan Belajar Fisika dengan hasil Belajar Fisika Siswa SMA Negeri 1 Duapitue Kabupaten Sidrap. *Jurnal Pendidikan Fisika UIN Alauddin Makassar*, 25–28. https://doi.org/10.24252/jpf.v2i1.3035

Lestari, A. D., Sadia, I. W., & Rachmawati, D. O. (2015). Pengaruh Kecemasan dan Motivasi Belajar terhadap Prestasi Belajar Fisika. Jurnal Pendidikan Fisika Undiksha, 2(1). https://doi.org/10.23887/jjpf.v2i1.5547

Linacre, J. M. (2002). Understanding Rasch measurement: Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*, 85–106.

Madoni, E. R., & Mardliyah, A. (2021). Determinasi Religiusitas, Kecerdasan Emosional, dan Dukungan Sosial Terhadap Kecemasan Akademik Siswa. *Jurnal Consulenza: Jurnal Bimbingan Konseling Dan Psikologi*, *4*(1), 1–10.

Molin, F., Cabus, S., Haelermans, C., & Groot, W. (2021). Toward Reducing Anxiety and Increasing Performance in Physics Education: Evidence from a Randomized Experiment. *Research in Science Education*, *51*(1), 233–249. https://doi.org/10.1007/s11165-019-9845-9

Muhaimin, Z. (2023). Pengaruh Pendidikan Agama Islam dalam Keluarga dan Budaya Religius Sekolah terhadap Kecerdasan Emosional Siswa SMA Negeri 2 Batu [Universitas Islam Negeri Maulana Malik Ibrahim Malang]. In *Etheses UIN Malang*. http://etheses.uin-malang.ac.id/id/eprint/53014

Mulyana, A. (2025). *Struktur Kurikulum Nasional Kurikulum Jenjang SD, SMP, SMA, dan SMK*. Forum Guru Nusantara. https://www.ainamulyana.info/

Musfah, J., Zakaria, R., Sofyan, A., Sayuti, W., Ridho, K., Fauzan, F., & Muawam, M. (2021). Pesantren-Based School Curriculum Integration Model in Indonesia. *MANAGERIA: Jurnal Manajemen Pendidikan Islam*. https://doi.org/10.14421/manageria.2020.52-13

Nazir, M. A., Izhar, F., Talal, A., Sohail, Z. B., Majeed, A., & Almas, K. (2021). A quantitative study of test anxiety and its influencing factors among medical and dental students. *Journal of Taibah University Medical Sciences*, *16*(2), 253–259. https://doi.org/10.1016/j.jtumed.2020.12.014

Pallant, J., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *The British Journal of Clinical Psychology / the British Psychological Society*, *46*, 1–18. https://doi.org/10.1348/014466506X96931

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Phafiandita, A. N., & Permadani, A. (2022). Urgensi Evaluasi Pembelajaran di Kelas. *Jurnal Inovasi Dan Riset Akademik*, *3*(2), 111–121.

Prasetyaningtyas, W. E., Rangka, I. B., Folastri, S., & Sofyan, A. (2022). Kecemasan Akademik Siswa di Sekolah: Suatu Tinjauan Singkat. *Journal of Learning and Instrucrional Studies*, *2*(3).

Purnama, Z. R., Laely, N., Hamidah, I. S., Ina, K., & Tukan, T. (2023). Hubungan Tingkat Kecemasan Belajar Dan Kemampuan Literasi Sains Dengan Hasil Belajar IPA Mahasiswa PGSD UNSAR. *Jurnal Administrasi Terapan*, *2*(2), 375–382. https://doi.org/10.31959/jat.v2i2.2056

Putranta, H., & Jumadi, J. (2019). Physics Teacher Efforts of Islamic High School in Yogyakarta to Minimize Students ' Anxiety When Facing the Assessment of Physics Learning Outcomes. *Journal for the Education of Gifted Young Scientists*, *7*(1), 119–136.

Rahmani, I., Kurniawan, E., & Sarjana, K. (2024). Pengaruh Kecemasan Matematika Terhadap Kemampuan Pemecahan Masalah Siswa Kelas XI IPA SMA Negeri 1 Gerung. *Jurnal Ilmiah Profesi Pendidikan*, *9*, 449–455.

Rahmawati, E., & Sari, I. M. (2025). Analisis Self efficacy Peserta Didik Dalam Pembelajaran Fisika Di Salah Satu SMA Kabupaten Cianjur. *JPF (Jurnal Pendidikan Fisika) FKIP UM Metro*, *13*(1), 1–13.

Regmi, P. R., Waithaka, E., Paudyal, A., Simkhada, P., & van Teijlingen, E. (2016). Guide to the design and application of online questionnaire surveys. *Nepal Journal of Epidemiology*, *6*(4), 640–644. https://doi.org/10.3126/nje.v6i4.17258

Revilla, L. (2009). Kecemasan Menghadapi Tes (Test Anxiety) Dan Dampaknya Terhadap Aktivitas Belajar. *Dinamika Ilmu: Journal of Education-Jurnal Pendidikan*, *9*. https://doi.org/10.21093/di.v9i2.287

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, *227*, 483–493. https://www.sciencedirect.com/science/article/pii/S0165032717303683

Wang, E. (2023). Physics: The Cornerstone of Scientific and Technological Development. *Journal of Macau University of Science and Technology (Humanities and Social Sciences)*, 1–6. https://doi.org/10.58664/mustjournal.2023.12.001

Zumbo, B. D. (1999). *A Handbook on the Theory and Methods of Differential Item Functioning (DIF)* (W. R. Wild (ed.)). Directorate of Human Resources Research and Evaluation National Defense Headquarters.