JoTaLP: Journal of Teaching and Learning Physics 10, 2 (2025): 131-145
Website: http://journal.uinsgd.ac.id/index.php/jtlp/index
ISSN 2580-3107 **(online) ISSN** 2528-5505 **(print)**

[Research Article]

EVALUATION OF HIGH SCHOOL TEACHERS' UNDERSTANDING OF FLIPPED CLASSROOM LEARNING USING THE RASCH APPROACH

Riskawati[1*], Khaeruddin[2], Nur Ifwa Awalya[3], Nur Aziza[4], Mifta Nurya Anwar[5], Wahyuni[6]

[123456]Department of Physics, Faculty of Mathematics and Natural Sciences, Makassar State University
Jl. Daeng Tata, Parang Tambung. Kec. Tamalate, Kota Makassar, Sulawesi Selatan 90224, Indonesia
E-mail: nurifwa2212@gmail.com

DOI: http://dx.doi.org/10.15575/jotalp.v10i2.45969
Received: 26 May 2025 ; Accepted: 25 August 2025 ; Published: 31 August 2025

ABSTRACT

The advancement of technology has influenced learning activities, including through the Flipped Classroom model, where students study independently at home and class time is used for discussion and collaboration. Teacher understanding and readiness are critical to its success. This study aims to analyze the level of understanding among elementary, junior, and senior high school teachers in South and West Sulawesi regarding the concept, challenges, and classroom implementation of the Flipped Classroom model. A descriptive quantitative method with a survey approach was applied, involving 49 randomly selected teachers. Data were analyzed using the Rasch model to identify the distribution and trends of teacher understanding. The results show that teacher understanding of the Flipped Classroom model is generally high and evenly distributed across demographic groups. These findings provide insight into designing more effective learning strategies across educational levels

Keywords: Teacher Understanding, Implementation, Learning Model, Flipped Classroom

How to cite: Riskawati, Khaeruddin, Awalya, N. I., Aziza, N., Anwar, M. N., & Wahyuni. (2025). Evaluation of high school teachers' understanding of flipped classroom learning using the Rasch approach, *Journal of Teaching and Learning Physics 10* (2), 131-145. DOI: http://dx.doi.org/10.15575/jotalp.v10i2.45969

1. INTRODUCTION

The flipped classroom has gained widespread attention as an innovative approach to increasing student engagement, facilitating independent learning, and optimizing learning time in the classroom. The flipped classroom allows teachers to dedicate more time in class to engaging and interactive learning activities that emphasize practice (Hastuti, 2020). The flipped classroom is a learning model in which students study the material at home in advance according to the tasks given by the teacher before learning in class (Agustini, 2021). The use of technology is one of the factors that supports the learning process, as students can easily access learning materials anytime and anywhere (Harianja, 2020). In this context, the Flipped Classroom not only maximizes time in class but also serves as a strategy to equip students with the skills needed in the 21st century.

Previous research conducted by (Mujiono, 2021) states that the use of the Flipped Classroom learning model provides benefits for both teachers and students. The time capacity provided by the Flipped Classroom learning model is able to accommodate a variety of students with different abilities and learning speeds. In line with this, research conducted by (Harianja, 2020) states that learning activities using the Flipped Classroom model provide good benefits in helping to increase students' curiosity and improve their learning outcomes in physics.

According to (Yasin & Judijanto, 2024), flipped classroom learning can make classroom learning more efficient and accommodate students with various backgrounds, characters, and abilities. Previous studies have also proven that the use of educational videos can increase the effectiveness of teaching and learning activities and save time (Sofianah & Sholihah, 2024). In addition, educators, parents, and students have a positive perspective on the implementation of the flipped classroom because it is considered to support the learning process (Sari & Hamami, 2022).

However, the Flipped Classroom learning model faces several challenges in its implementation, particularly in the city of Makassar. The main challenge is the readiness of teachers and students to adapt technology as part of the learning process (Ramadhani et al., 2022). Not all students have adequate access to devices and stable internet connections, which can hinder the effectiveness of independent learning at home (Kardika et al., 2023). Not all students can be motivated to study independently at home. (Wijayanto et al., 2022). Teachers also need to adjust their teaching methods to be more interactive and able to encourage student involvement in collaborative activities in the classroom (Sri Lestari, 2021).

Among the various challenges that exist, this model can be an innovative strategy that can improve the effectiveness of the teaching and learning process in the digital age (Amelia et al., 2021). To implement this model optimally, teachers need to be adaptable to technology, creative in developing digital learning materials, able to guide students to be more independent in their learning process (Santoso, 2025), and most importantly, in line with Sardiman's opinion that learning outcomes will be optimal if there is motivation. The more appropriate the motivation provided, the more successful the lesson will be. Thus, motivation determines the intensity of students' learning efforts (Sharfina & Soraya, 2024).

Therefore, this study aims to describe the level of teachers' understanding of the Flipped Classroom learning model, identify the challenges and obstacles faced by teachers in its implementation, and evaluate the application of this model in elementary, junior high, and high schools in South Sulawesi and West Sulawesi provinces.

2. METHOD

2.1 Participants

This study is a quantitative descriptive study using a survey approach to determine the level of understanding of elementary, junior high, and high school teachers in South Sulawesi and West Sulawesi provinces regarding the Flipped Classroom learning model. This approach was chosen to obtain a descriptive overview of the competence and application of the Flipped Classroom in learning activities at various levels of education. The data collection instrument was a questionnaire developed based on the concept of the Flipped Classroom, covering aspects of basic concept understanding, advantages and disadvantages, and application in the classroom. The questionnaire used a four-point Likert scale from 1 point (strongly disagree) to 4 (strongly agree) to measure teachers' perceptions and understanding quantitatively.

The collected data were analyzed using the Rasch measurement model, which allows for objective assessment of respondents' latent abilities as well as the validity and reliability of the questions in the instrument. This approach provides an accurate picture of the distribution of teachers' understanding of Flipped Classroom learning. A total of 49 teachers from elementary, junior high, and high schools spread across South Sulawesi and West Sulawesi participated in this study. Data collection was conducted online using Google Forms to reach participants from various backgrounds. The demographic information collected included education level, gender, teaching location, and teaching interests. This demographic data is important for understanding the context of Flipped Classroom implementation and the factors that influence teachers' level of understanding of the Flipped Classroom learning model.

Table 1. Demographic Profile of Participants

Aspect	Category	f	(%)
Gender	Female	32	65,31
	Male	17	34,69
Level of Education	Elementary School	45	91,88
	Junior high school	3	6,12
	High school	1	2
Age	20-29 y.o	11	22,45
	30-39 y.o	26	53,1
	40-49 y.o	11	22,45
	> 50 y.o	1	2
Lesson	Non-science	28	57,14
	Science	21	42,86

2.2 Instrument

The questionnaire in this study was developed to measure teachers' understanding of the Flipped Classroom learning model. The instrument consists of 25 items covering the basic concepts, objectives, benefits, stages, characteristics, readiness, application, and obstacles and solutions in Flipped Classroom learning. Each item uses a four-point Likert scale, from 1 (Very Unfamiliar) to 4 (Very Familiar). The questionnaire is divided into two parts. The first part contains background data on respondents, such as age, gender, subject taught, and teaching level. The second part contains statements based on seven main indicators, namely Understanding of the Basic Concepts of Flipped Classroom, Objectives and Benefits of Flipped Classroom Learning, Stages of the Flipped Classroom Learning Model, Characteristics of the Flipped Classroom Learning Model, Teacher Readiness in Implementing Flipped Classroom, Application of Flipped Classroom Learning, and Obstacles and Solutions in the Implementation of Flipped Classroom. The collected data was inputted and analyzed using Ministep software based on the Rasch model. The following table shows the indicators and questionnaire items in the Flipped Classroom.

Tabel 2. Indikator dan Item kuesioner dalam *Flipped Classroom*

Indikator	Nomor Item
Pemahaman Konsep Dasar Model Pembelajaran Flipped Classroom	1, 2, 3
Pemahaman Tujuan dan Manfaat Pembelajaran Flipped Classroom	4, 5, 6
Pemahaman Tahapan Model Pembelajaran Flipped Classroom	7, 8, 9, 10, 11
Karakteristik Model Pembelajaran Flipped Classroom	12, 13, 14, 15
Kesiapan Guru dalam Mengimplementasikan Flipped Classroom	16, 17, 18
Penerapan Pembelajaran Flipped Classroom di Kelas	19, 20, 21
Kendala dan Solusi dalam Implementasi Flipped Classroom	22, 23, 24, 25

2.3 Procedure

Data collection was conducted for approximately two months using the Google Forms platform before the final analysis was carried out. The Flipped Classroom learning model questionnaire, which included a main scale and demographic items, was distributed digitally through academic forums and social media platforms. The target population consisted of teachers teaching at elementary, junior high, and high schools throughout South Sulawesi and West Sulawesi. Participants were recruited voluntarily through online distribution strategies, including educational WhatsApp groups and digital learning communities.

To ensure that participants understand the completion procedure, clear instructions are provided at the beginning of the online questionnaire. Participants are informed that all items must be answered without omission to maintain data completeness and response quality. The Google Form system automatically records responses once they are submitted. Confidentiality and anonymity are strictly maintained by assigning codes to responses and not collecting personal information. Data integrity is further ensured through completeness checks, with only fully completed responses included in the final dataset.

2.4 Data Analysis

Perangkat lunak MINISTEP versi 5.7.1 (Linacre, 1991) dimanfaatkan dalam analisis Rasch dengan metode Joint Maximum Likelihood Estimation (JMLE) guna mengevaluasi karakteristik psikometrik serta skala penilaian yang digunakan. Pendekatan Rasch dalam penelitian ini berfungsi untuk mengubah data respons ordinal siswa menjadi data interval dalam satuan logit, yang memiliki rentang nilai dari negatif hingga positif tanpa batas. Untuk menilai kualitas instrumen Kuesioner *Flipped Classroom* dilakukan analisis terhadap sejumlah indikator Rasch, antara lain unidimensionalitas, independensi lokal, pemisahan antara item dan responden, reliabilitas item, serta validitas berdasarkan kecocokan nilai.

Mean Square (MNSQ) infit and outfit. Unidimensionality is considered to be satisfied if the percentage of raw variance explained by the measurement exceeds 30%, and the unexplained variance of the first contrast is below two eigen values (Boone et al., 2014). To test for local independence, Yen's Q3 statistic is used, with a criterion of raw residual correlation less than 0.4 (Yen, 1993). Item and respondent separation values exceeding 2 logits indicate the presence of significantly different groups in the data (Bond & Fox, 2015). Internal consistency and measurement stability are considered adequate if item reliability and Cronbach's alpha (α) values exceed 0.60 (DeVellis, 2017; Tavakol & Dennick, 2011). Convergent validity is evaluated through the infit and outfit MNSQ values, with an ideal range between 0.5 and 1.5; however, values up to 1.6 are *still considered acceptable as long as the Point-Measure Correlation (PTMA) value remains positive (Wright et al., 1994).*

Wright's map, which illustrates the distribution between respondent ability and item difficulty, was used to evaluate the suitability of the interaction between the two. To confirm the suitability of items at the instrument level, the Item Characteristic Curve (ICC) was used. In addition, Item Differential Function (DIF) analysis was conducted based on gender to identify potential bias in the items. Finally, a logit size histogram visualization was created using R software to visually compare the level of understanding of the Flipped Classroom between male and female participants.

3. RESULT AND DISCUSSION

3.1 Validity

The validation of the instrument for measuring teachers' understanding of the Flipped Classroom learning model was conducted using Rasch analysis with the Joint Maximum Likelihood Estimation (JMLE) approach. This analysis was used to assess the psychometric strength of the developed questionnaire to ensure that the instrument could measure teachers' understanding reliably and objectively. The assessment of the suitability between respondents and items was carried out through the Mean Square infit and outfit values (MNSQ), with an acceptance limit ranging from 0.5 to 1.5. For large sample sizes, values up to 1.6 are still tolerable (Boone et al., 2014; Linacre, 1991).

In this study, all items and respondents showed MNSQ values within the ideal range, indicating that teachers' responses to each item were in line with the Rasch model expectations. This indicates that the instrument has strong internal validity and can consistently distinguish between levels of teacher understanding. Although Z-standard values (ZSTD) were also generated in the Rasch analysis, these statistics were not the main focus of this study due to their high sensitivity to large sample sizes. In this study, the values were above the minimum threshold of 2.0 logits, indicating that the difficulty levels of the items varied considerably and could effectively distinguish between teachers with low and high levels of understanding. In addition, the reliability of the items was close to the maximum value of 1.00, indicating a high degree of consistency in item calibration and ensuring that this instrument can be reused in similar measurements in the future (Boone et al., 2014).

The unidimensionality aspect was also evaluated based on the proportion of raw variance, which exceeded 30% for each subscale. These results indicate that each part of the questionnaire measures the same construct, namely teachers' understanding of the Flipped Classroom learning model. In addition, the unexplained variance in the first contrast is below an eigen value of 2.0, indicating that there are no significant secondary dimensions (Bond & Fox, 2015; Wright et al., 1994).

Table 3. Summary of Rasch Parameters for the Flipped Classroom Model Instrument and Each Subscale

Psychometric Attribute	A	B	C	D	E	F	G	Teachers' Understanding
Number of Items	3	3	5	4	3	3	4	25
Mean								
Item Outfit Mnsq	1.03	0.88	2.30	0.58	1.21	1.17	0.54	1.15
Item Infit Mnsq	0.97	0.96	0.90	0.97	0.87	0.88	0.98	0.97
Person Outfit Mnsq	1.03	0.88	1.03	0.58	1.02	0.94	0.54	1.15
Person Infit Mnsq	0.98	0.85	0.66	0.49	0.68	0.86	0.54	0.92

Item Separation	0.23	1.83	3.22	1.16	3.62	4.01	0.63	2.73
Person Separation	0.00	0.46	1.29	0.00	0.63	0.79	0.00	3.42
Item Reliability	0.05	0.77	0.91	0.57	0.93	0.94	0.11	0.88
Cronbach's Alpha	0.57	0.68	0.83	0.89	0.73	0.77	0.85	0.96
Unidimensionality								
Raw Variance Explained by Measure (%)	16.4	74.6	62.5	61.2	62.8	66.5	41.4	52.5
Unexplained Variance 1st Contrast	20.188	16.243	18.740	20.172	15.700	19.074	15.484	36.227

Table 3 presents a summary of Rasch parameters for instruments in the Flipped Classroom learning model, which includes seven main indicators. In general, the psychometric characteristics of these instruments show adequate quality for students in the context of Flipped Classroom-based learning.

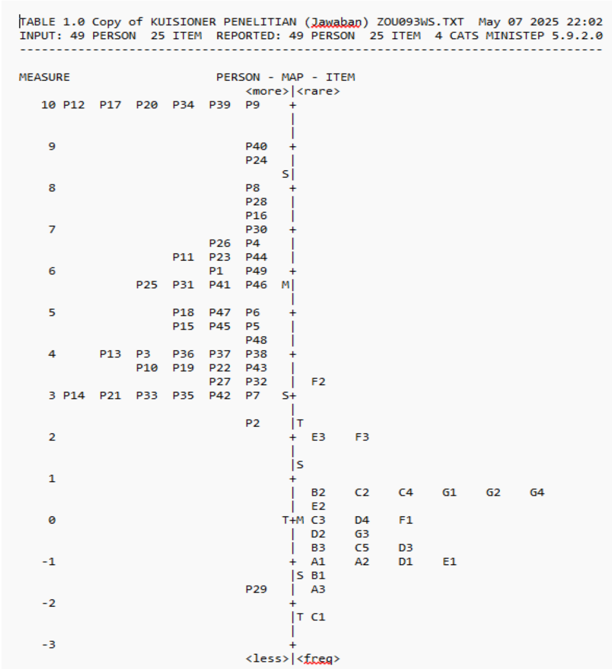
The highest item separation value was recorded on indicator F (4.01), which shows the instrument's ability to distinguish the level of difficulty between items on that indicator. Meanwhile, the highest person separation value was found on indicator G (0.90) and the questionnaire as a whole (3.42), indicating that respondents could be grouped into three or more different ability levels consistently.

The highest item reliability was achieved by indicators F (0.94), G (0.91), and the overall scale (0.88), indicating that the items on these indicators were very stable and reliable. Conversely, indicator A showed a very low item reliability value (0.05), indicating the need to revise the items on this indicator so that it can differentiate respondents more effectively. Cronbach's Alpha overall was 0.96, reflecting excellent internal consistency between items on the scale (DeVellis, 2017; Tavakol & Dennick, 2011). The highest Alpha value was achieved on indicator G (0.96), while the lowest value was 0.57 on indicator A.

The MNSQ values for outfit and infit are also within the recommended range (0.5–1.5) according to Linacre (2022), both for items and participants. On the overall scale, the MNSQ outfit item value is 1.15 and the MNSQ infit item value is 0.97. For respondents, the MNSQ outfit value is 1.15 and the MNSQ infit is 0.92. This confirms that there are no extreme deviations between the respondents' answer patterns and the model predictions, thus supporting the construct validity of the instrument (Boone et al., 2014).

From a unidimensionality perspective, the percentage of raw variance explained by the measure for the overall scale is 52.5%, which exceeds the minimum threshold of 30% and indicates that most of the variance in the data is explained by one main dimension, namely the AI literacy construct in the context of Flipped Classroom learning. In addition, the unexplained variance value in the first contrast of 3.6227 is still within the tolerance limit, although it indicates the possible existence of minor subdimensions (Bond & Fox, 2015), such as teachers' understanding of the Flipped Classroom..

Several indicators, such as A, require review of their items, but in general, the instrument has met the validity criteria based on the Rasch model. This instrument can be used to accurately and consistently evaluate teachers' understanding of literacy in Flipped Classroom learning.



```
TABLE 1.0 Copy of KUISIONER PENELITIAN (Jawaban) ZOU093WS.TXT  May 07 2025 22:02
INPUT: 49 PERSON  25 ITEM  REPORTED: 49 PERSON  25 ITEM  4 CATS MINISTEP 5.9.2.0
--------------------------------------------------------------------------------

MEASURE                   PERSON - MAP - ITEM
                                <more>|<rare>
   10 P12  P17  P20  P34  P39  P9     +
                                      |
                                      |
    9                          P40    +
                               P24    |
                                     S|
    8                          P8     +
                               P28    |
                               P16    |
    7                          P30    +
                          P26  P4     |
                     P11  P23  P44    |
    6                     P1   P49    +
                P25  P31  P41  P46   M|
                                      |
    5                P18  P47  P6     +
                     P15  P45  P5     |
                               P48    |
    4      P13  P3   P36  P37  P38    +
                P10  P19  P22  P43    |
                          P27  P32    |  F2
    3 P14  P21  P33  P35  P42  P7    S+
                                      |
                               P2     |T
    2                                 +  E3    F3
                                      |
                                      |S
    1                                 +
                                      |  B2    C2    C4    G1    G2    G4
                                      |  E2
    0                                T+M C3    D4    F1
                                      |  D2    G3
                                      |  B3    C5    D3
   -1                                 +  A1    A2    D1    E1
                                      |S B1
                               P29    |  A3
   -2                                 +
                                      |T C1
                                      |
   -3                                 +
                                <less>|<freq>
```

Figure 1. *Wright map*

The Wright Map in Figure 1 above provides a clear visualization of the relationship between teachers' level of understanding of the Flipped Classroom learning model and the level of difficulty of items in the questionnaire developed. On the left side of the map, the distribution of respondents' abilities (elementary, junior high, and high school teachers) is displayed vertically in logit units, with teachers who have a higher level of understanding at the top, while those with lower understanding are at the bottom. The right side shows the level of difficulty of the items on the questionnaire, arranged from the most difficult (at the top) to the easiest (at the bottom). The relatively symmetrical distribution around the average logit value of 0.00 indicates a balance between respondent ability and item difficulty, which signifies good targeting quality of the measurement tool, as recommended in the Rasch measurement model.

The distribution of teacher abilities is quite spread out, with respondents concentrated in the logit range of approximately +1 to +4, marked by an accumulation of # symbols. This indicates that most teachers have a medium to high level of understanding. On the other hand, items such as F2 (highest logit), F3, and E3 are at the top logit position, indicating that these items measure more complex aspects of understanding the Flipped Classroom learning model. Referring to the further evaluation, these three items need to be reviewed, especially if they show MNSQ Infit values that are far above the acceptable limit. Conversely, items such as C1 (logit = –3), B3, and C2 are at the lowest logit, reflecting items that are relatively easy and measure more basic understanding. These items can be categorized as items with a low level of difficulty and need to be considered whether they still provide sufficient information to distinguish the abilities of respondents.

Overall, although the MNSQ Infit and ZSTD Outfit values are not presented in this Wright Map, most items appear to be proportionally distributed across the respondents' ability distribution. This indicates that many items are likely to function as expected psychometrically (Bulala et al., 2024), (Ehrich et al., 2021). However, there are several items at the upper and lower ends of the spectrum (such as F2, F3, and C1, B3) that may need to be revised or reanalyzed to ensure that their content and level of interpretation are consistent with the construct of teacher understanding being measured.

3.2 Reliability

Reliability criteria were evaluated using several indicators, including Rasch parameters through respondent reliability (person reliability) and item reliability (Fisher, 2007; Linacre, 2022), as well as Cronbach's Alpha (α) (Taber, 2018). The results of the analysis show that the item reliability values for each indicator range from 0.00 to 0.94. The highest value is found in indicator F (0.94), while the lowest value is found in indicator A (0.00), indicating that some items still need to be reviewed to improve measurement stability. Meanwhile, person reliability ranged from 0.00 to 0.92. The highest value was achieved by indicator G (0.92), while the lowest value was 0.00 in indicators A and F. This reflects a significant difference in the respondents' ability to answer certain items, as well as the possibility that some instruments need to be improved in order to more accurately distinguish the respondents' level of understanding.

The Cronbach's alpha (α) value indicates the level of internal consistency of the instrument, with a range between 0.57 and 0.85. The highest value was obtained by indicator G (0.85), and the lowest value by indicator A (0.57). According to Taber's (2018) criteria, a value above 0.70 indicates acceptable reliability, so most indicators in this instrument show fairly good to excellent internal consistency. All reliability results for each indicator are summarized in Table 4.

Table 4. Reliability Indicators

Reliability	Instrumen							Pembelajaran Flipped Classroom
	A	B	C	D	E	F	G	
Item reliability	0.00	0.77	0.83	0.57	0.93	0.94	0.11	0.88
Person reliability	0.05	0.17	0.62	0.00	0.28	0.38	0.00	0.92
Cronbach's alpha (α)	0.57	0.69	0.68	0.89	0.73	0.77	0.85	—

3.3 DIF Based on Gender and Age

Differential Item Functioning (DIF) analysis in this study was conducted to evaluate whether there were items in the Flipped Classroom-based learning model comprehension instrument that showed potential bias toward certain respondent groups, particularly based on gender and age. DIF evaluation is important in the Rasch context to ensure that items in the instrument are fair and free from demographic bias. The DIF classification criteria refer to standard guidelines: negligible (|DIF| < 0.43 logit), mild to moderate (|DIF| ≥ 0.43), and moderate to large (|DIF| ≥ 0.64) (Boone et al., 2014; Zwick, 2012). Statistically significant DIF logit values ($p < 0.05$) form the basis for determining whether an item shows bias toward a particular group.

Based on the results of DIF analysis by gender (Figure 2), it can be seen that several items have differences in logit values between male (M) and female (F) respondents. Three items show DIF logit values that exceed the threshold of 0.43 logit, indicating mild to moderate bias. The pattern in the graph shows a tendency for items on the right side (with positive logits) to favor the male group, while other items show negligible differences. Overall, only a small number of items show potential gender bias, while most items remain invariant, meaning that the instrument can still be used fairly by both female and male teachers.

Furthermore, analysis based on age groups (Figure 3) revealed that most items had DIF values in the low range and were not practically significant. Respondents were divided into four age groups: 20–29, 30–39, 40–49, and >50 years. Although there were logit fluctuations between groups, the highest logit values were recorded in the 20–29 age group, followed by the 30–39, >50, and 40–49 age groups. Several items, as shown on the right side of the graph, appear to have logit values slightly approaching the 0.43 logit threshold, but none of the items exceed the 0.64 logit threshold, which indicates high bias. For example, one item from indicator F appears to touch the logit threshold of ±0.48, but it is still in the category of mild bias.

Overall, the results of the DIF analysis based on both gender and age show that the Flipped Classroom instrument sufficiently meets the principles of fairness and invariance, which are key requirements in Rasch-based psychometric measurement. These results are also in line with previous literature that emphasizes the importance of evaluating demographic bias in instrument validation (Soeharto et al., 2024). Thus, although there are several items that show a slight bias, this instrument is still considered valid and reliable for use by teachers of various age and gender backgrounds in the context of Flipped Classroom-based learning.

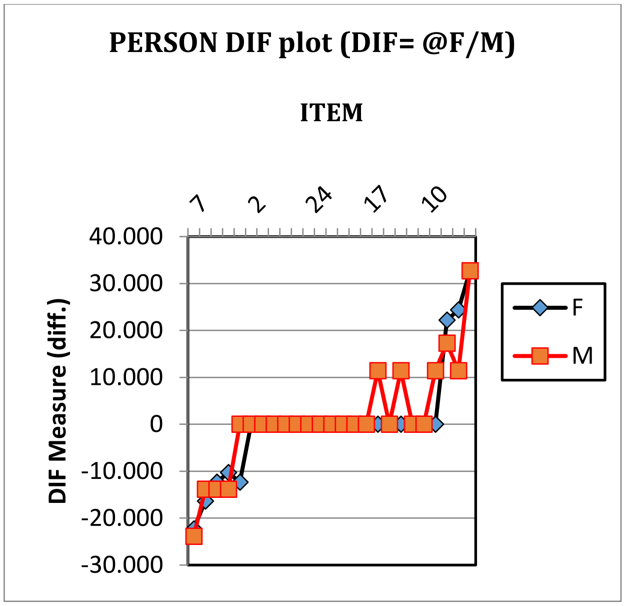


Figure 2. DIF Analysis Based on Gender

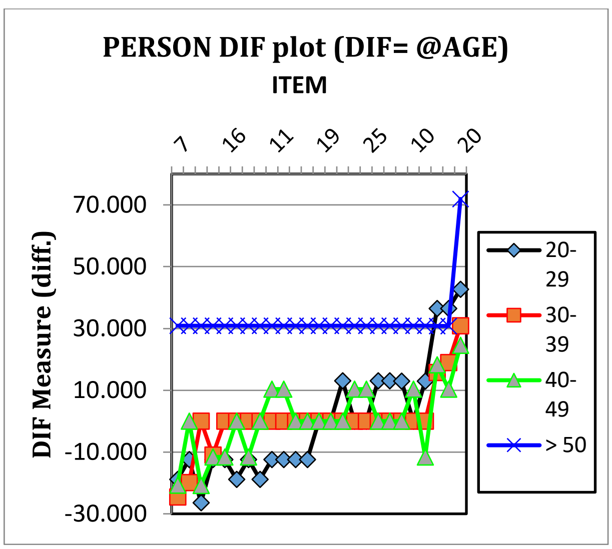


Figure 3. DIF Analysis Based on Age

3.4 DIF Based on Teaching Level and Subject Taught

Differential Item Functioning (DIF) analysis was conducted to identify potential inequalities in measuring teacher literacy in the Flipped Classroom learning model, particularly in terms of teaching levels (elementary, junior high, and high school) and subject areas (science and non-science). This evaluation is important to ensure that each item in the instrument functions fairly across all groups of teachers without showing bias toward certain demographic characteristics. Based on the DIF results at the teaching level (Figure 4), most items show DIF values within the low tolerance range (|DIF| < 0.43 logit), indicating that these items function relatively consistently for elementary, junior high, and high school teachers. However, there are noticeable deviations in certain items, such as item 23, where the DIF score for high school teachers is much higher than that for elementary and junior high school teachers. This indicates a potential bias towards high school teachers, possibly due to their greater experience and exposure to the Flipped Classroom practice, which is generally more widely applied at the senior secondary level. This tendency should be noted so as not to neglect the needs of teachers at the elementary and junior high school levels in training and evaluation of digital literacy competencies.

Furthermore, the DIF graph based on subject area (Figure 5) shows an almost parallel response pattern between science and non-science teachers. All items are within the low DIF range, and none exceed the logit threshold of ≥ 0.43. This indicates that subject area background does not influence how teachers understand and respond to items on the Flipped Classroom instrument. In other words, both science and non-science teachers demonstrate similar levels of interpretation and understanding of the concepts measured in the instrument

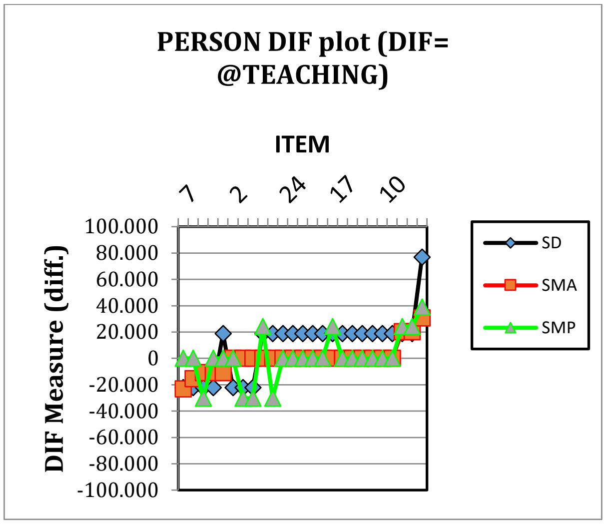


Figure 4. DIF Analysis Based on Teaching Level

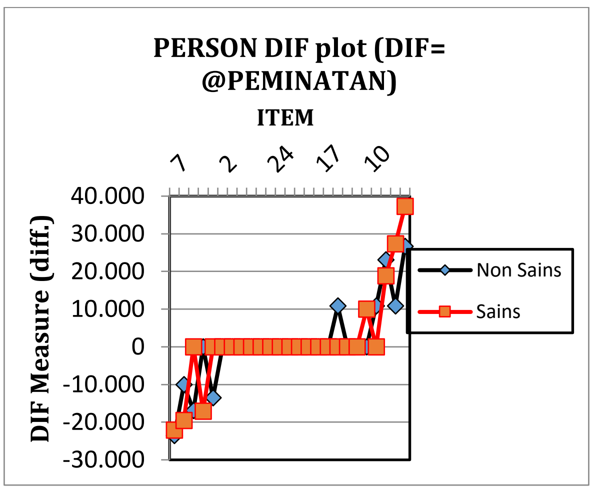


Figure 5. DIF Analysis Based on Subject Taught

3.5 Evaluation of Elementary, Junior High, and Senior High School Teachers Understanding of the Flipped Classroom Model in West Sulawesi Province

Table 5. Teachers' Understanding of the Flipped Classroom Model in Elementary, Junior High, and Senior High Schools Based on Age, Gender, Teaching Level, and Subject Area

Demographic Group	N	M (score)	M (logit)	SD	Person Reliability
Gender					
Female	32	82.3	0.43	2.41	0.93
Male	17	88.5	0.70	2.81	0.88
Age					
20–29 years	11	86.5	0.79	2.72	0.92
30–39 years	26	81.5	0.45	2.24	0.90
40–49 years	11	88.4	0.95	3.10	0.92
> 50 years	1	99.0	–	0.00	0.90
Educational Level					
Elementary School (SD)	28	84.0	0.53	2.73	0.93
Junior High School (SMP)	21	85.2	0.60	2.69	0.91
Senior High School (SMA)	1	81.0	–	0.00	0.00
Subject Taught					
Science	3	82.7	1.86	2.63	0.94
Non-science	45	84.7	0.41	2.75	0.92

Analysis of teachers' understanding scores of the Flipped Classroom learning model based on gender, age, education level, and subject type shows demographic trends that reflect variations in teacher competency achievement (Table 5). Male teachers obtained a higher average logit score for understanding (M = 0.70; SD = 2.81) than female teachers (M = 0.43; SD = 2.41). However, the level of person reliability among female teachers was slightly higher (0.93) than that among male teachers (0.88), which still shows the consistency of the instrument in measuring understanding of the Flipped Classroom based on the Rasch model. These findings are consistent with Pristiani et al. (2023), who stated that differences in understanding can be influenced by socio-educational factors and access to technology.

Based on age group, teachers aged >50 years showed the highest logit score (M = 1.00; SD = 0.00), followed by the 40–49 age group (M = 0.95; SD = 3.10), then 20–29 (M = 0.79; SD = 2.72), and 30–39 (M = 0.45; SD = 2.24). This pattern shows that senior teachers tend to have a higher understanding of the Flipped Classroom, possibly due to their longer pedagogical experience, although the person reliability in this group is relatively uniform.

In terms of education level, junior high school teachers showed the highest logit score for understanding (M = 0.60; SD = 2.69), followed by elementary school teachers (M = 0.53; SD = 2.73). Meanwhile, high school teachers showed significantly lower scores (M = 0.00; SD = 0.00), but this figure appears to be an outlier or anomalous data because it is far from the normal logit range in Rasch measurements.

From the perspective of the subjects taught, science teachers showed higher average logit scores (M = 1.86; SD = 2.63) than non-science teachers (M = 0.41; SD = 2.75), with better person reliability (0.94 vs. 0.92). This indicates that Flipped Classroom integration occurs more frequently in exact or laboratory-based disciplines, which require intensive use of technological tools (Chaw & Tang, 2023; Kubota et al., 2025).

Overall, the relatively high reliability (average of approximately 0.90) across almost all demographic groups indicates that the instrument used has a good level of stability and precision.

Table 6. Categorization of Item Difficulty Levels of the Flipped Classroom Learning Model

Task	Difficulty level I, LVI ≥ Mean logit + 2SD	Difficulty level II, Mean logit + 2SD > LVI ≥ 1SD	Difficulty level III, 1SD > LVI ≥ Mean logit	Difficulty level IV, Mean logit > LVI ≥ -1SD	Difficulty level V, LVI < -1SD
A				A1 A2	A3
B			B2	B3	B1
C			C3 C2 C4	C5	C1
D			D4	D2 D3 D1	
E		E3	E2	E1	
F	F2	F3	F1		
G			G4 G1 G2	G3	

Item Level Analysis (ILA) is used to measure the level of difficulty of each item in the questionnaire on teachers' understanding of the Flipped Classroom learning model, using the Rasch approach. The logit reference point is set at a mean of 0.00 with a standard deviation (SD = 1.12), so that the classification of item difficulty is divided into five categories: very difficult (LVI ≥ +2SD), difficult (+2SD > LVI ≥ +1SD), moderate (+1SD > LVI ≥ mean), easy (mean > LVI ≥ −1SD), and very easy (LVI < −1SD). This approach provides a detailed picture of the distribution of teachers' understanding of the Flipped Classroom learning model as reflected in the research instrument.

Based on the analysis results, most items in the instrument fall into the low difficulty category (categories IV and V). There are no items in the very difficult category (category I), indicating that all statements are relatively well understood by respondents. In the difficult category (category II), there is only one item, namely G3, which indicates that this item is still a challenge for some teachers and requires further attention. In the moderate difficulty category (category III), there are several items from groups C (C3, C2, C4, C5), D (D4), E (E2, E3), and F (F3).

The easy category (category IV) covers most of the items in groups A (A1, A2), B (B2, B3), D (D2, D3), and F (F1), indicating that the majority of teachers have a good understanding of the elements of the Flipped Classroom represented by these groups. Meanwhile, the very easy category (category V) includes items A3, B1, and C1, which indicate that these items are very easy for respondents to understand. Meanwhile, group G is the only group that has items in the high difficulty category, with G3 included in the difficult category (category II) and G4, G1, G2 in the moderate category (category III). This indicates that the aspects measured by group G tend to be more complex and still pose a challenge for most teachers.

Table 7. Categorization of Teachers Based on the Flipped Classroom Learning Model

Demographic Group	N	M (score)	M (logit)	SD	Person Reliability
Gender					
Female	32	82.3	0.43	2.41	0.93
Male	17	88.5	0.70	2.81	0.88
Age					
20–29 years	11	86.5	0.79	2.72	0.92
30–39 years	26	81.5	0.45	2.24	0.90
40–49 years	11	88.4	0.95	3.10	0.92
> 50 years	1	99.0	–	0.00	0.90
Teaching Level					
Elementary (SD)	28	84.0	0.53	2.73	0.93
Junior High (SMP)	21	85.2	0.60	2.69	0.91
Senior High (SMA)	1	81.0	–	0.00	0.00
Subject Area					
Science	3	82.7	1.86	2.63	0.94
Non-science	45	84.7	0.41	2.75	0.92

Analysis of the Learner Variable Profile (LVP) in the context of teachers' understanding of the Flipped Classroom learning model was conducted using a quantitative approach with Rasch analysis. The classification of understanding levels was based on personal logit scores, with understanding categories divided into four levels: very high (LVP ≥ mean logit + 2SD), high (mean logit + 2SD > LVP ≥ mean logit), moderate (mean logit > LVP ≥ mean logit − 2SD), and low (LVP < mean logit − 2SD). This classification became the basis for assessing the distribution of teachers' understanding of the Flipped Classroom learning model based on demographic variables. In terms of gender, the majority of respondents were female teachers (75%), who were dominated by the high (30 teachers) and very high (1 teacher) understanding categories. Male teachers, although fewer in number, also showed a good distribution, with 14 teachers in the high category and 3 teachers in the very high category. There were no teachers, either female or male, in the low or very low categories. This shows that both female and male teachers generally have a good understanding of the concept and implementation of the Flipped Classroom, with male teachers tending to have a higher proportion in the very high category.

In terms of age, teachers aged 30–39 years old were the largest group that showed a high understanding of the Flipped Classroom model, with 24 people in the high category and 1 person in the very high category. Teachers aged 20–29 years old contributed 10 people in the high category and 1 person in the very high category. Teachers aged 40–49 were all in the high category (9 people), and only one teacher over the age of 50 was in the high category. These data show that teachers from various age groups, especially the productive age group (20–39 years), are highly prepared to adopt the Flipped Classroom learning model, with a good adaptive tendency towards new approaches to learning.

In terms of education level, all respondents were at the high school level (40 people), junior high school level (3 people), and elementary school level (1 person). All high school teachers had a high level of understanding (40 people), while all junior high school teachers were in the high category (3 people), and elementary school teachers were in the very high category (1 person). This indicates that high school teachers dominate in terms of the number and consistency of their understanding of the Flipped Classroom. This tendency can be attributed to exposure to modern learning strategies and the availability of more complete supporting facilities at the high school level.

Viewed from the subjects taught, non-science teachers were more dominant in terms of numbers, but science teachers showed a higher proportion in the very high category. Of the total 27 non-science teachers, 25 were in the high category and 2 in the very high category. Meanwhile, there were 17 science teachers, all of whom were in the high (15) and very high (2) categories, with none in the medium or low categories. These findings indicate that both science and non-science teachers possess a high level of understanding of the Flipped Classroom model, although science teachers appear to be slightly superior in achieving very high levels.

Overall, the results of this LVP analysis show that teachers' understanding of the Flipped Classroom learning model is generally at a high level, with an even distribution across various demographic groups. No teachers were found to be in the low category, indicating a generally good level of readiness to implement this model. These findings provide an important basis for designing teacher training and professional development programs so that the Flipped Classroom strategy can be implemented more optimally at various levels and fields of education.

The results of this study indicate that the instrument used to measure teachers' understanding of the Flipped Classroom learning model has good psychometric quality, in accordance with the Rasch approach. Item reliability values reaching 0.88 and respondent reliability of 0.92 reflect high internal consistency, as well as the instrument's ability to effectively differentiate teachers' levels of understanding. All outfit and infit Mean Square (MNSQ) values were within the expected range of 0.5–1.5, indicating the suitability of the data for the Rasch model (Linacre, 2022; Boone et al., 2014). The resulting Wright Map showed a balanced distribution between teacher ability and item difficulty. The unidimensionality of the instrument was also quite good, with raw variance explained by measure reaching 52.5%, which exceeded the minimum threshold of 30% (Bond & Fox, 2015). Unexplained variance in the 1st contrast was below the eigen value of 2.0, indicating that there were no significant secondary dimensions. At the subscale level, raw variance reached more than 60%, which strengthened the internal structure of the instrument. Differential Item Functioning (DIF) analysis showed that there was no significant bias based on gender, age, education level, or subject area.

4. CONCLUSION

The results of this study indicate that teachers' understanding of the Flipped Classroom learning model is generally high, with an even distribution across various demographic groups, such as gender, age, education level, and subject area. There were no respondents in the low understanding category, indicating that teachers are ready to implement this model in their respective schools. The instruments used in this study met good psychometric characteristics.

Analysis using the Rasch model showed high validity and reliability, both at the item level and at the respondent level. The Wright Map results showed a balance between the teachers' ability level and the item difficulty level, which supports the effectiveness of the instrument in evaluating teachers' understanding of the Flipped Classroom. However, this study also revealed a number of challenges that teachers still face in applying this model, particularly related to limited access to technology, students' readiness for independent learning, and teachers' ability to design interactive learning. Several items also showed a high level of difficulty and need further attention in the development of future teacher training. In general, these findings reinforce that the Flipped Classroom has great potential to be implemented more widely in the education system.

5. REFERENCES

Agustini, M. (2021). Meningkatkan Aktivitas Dan Hasil Belajar Siswa Menggunakan Model Flipped Classroom melalui Aplikasi Google Classroom. *Indonesian Journal of Educational Development*, *2*(2), 280–289. https://doi.org/10.5281/zenodo.5244420

Amelia, A., Nurfalah, E., & Mulhayatiah, D. (2021). Peningkatan Tpack Guru Fisika Melalui Media Pembelajaran Berbasis Google Apps. *Journal of Teaching and Learning Physics*, *6*(2), 69–76. https://doi.org/10.15575/jotalp.v6i2.10298

Bond, T., & Fox, C. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (T. Bond, Ed.; 3rd ed.). Routledge.

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

Bulala, T., Anane, E., & Cobbinah, A. (2024). Psychometric Assessment of Botswana's 2018 Junior Certificate Agriculture Science Items through Item Response Theory. *EAST AFRICAN JOURNAL OF EDUCATION AND SOCIAL SCIENCES*, *5*(1), 19–26. https://doi.org/10.46606/eajess2024v05i01.0345

DeVellis, R. F. (2017). *Scale Development Theory and Applications* (A. Villaruel, Ed.; 4th ed.). SAGE.

Ehrich, J. F., Howard, S. J., Bokosmaty, S., & Woodcock, S. (2021). An Item Response Modeling Approach to Cognitive Load Measurement. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.648324

Harianja, J. K. (2020). Mengembangkan Sikap Rasa Ingin Tahu (Curiosity) Pada Pelajaran Fisika Menggunakan Model Pembelajaran Flipped Classroom. *Jurnal Pendidikan Fisika Dan Teknologi*, *6*(1), 121–130. https://doi.org/10.29303/jpft.v6i1.1738

Hastuti, W. D. (2020). Membangun Motivasi dan Kemandirian Peserta Didik Berkebutuhan Khusus Melalui Flipped Classroom di Masa New Normal Covid-19. *Prosiding Webinar Magister Pendidikan Nonformal UNG | 181*, *September*, 181–192.

Kardika, R. W., Rokhman, F., & Pristiwati, R. (2023). Penggunaan Media Digital terhadap Kemampuan Literasi Multimodal dalam Pembelajaran Bahasa Indonesia. *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, *6*(9), 6715–6721. https://doi.org/10.54371/jiip.v6i9.2307

Linacre, J. M. (1991). *A user's guide to WINSTEPS ministep : rasch-model computer programs*. Winsteps.com.

Mujiono, N. (2021). Flipped Classroom: Sekolah Tanpa Pekerjaan Rumah. *Jurnal Teknodik*, *25*(1), 67. https://doi.org/10.32550/teknodik.v25i1.457

Ramadhani, R., Syahputra, E., & Simamora, E. (2022). Ethno-Flipped Classroom Model: Sebuah Rekomendasi Model Pembelajaran Matematika Di Masa New Normal. *AXIOM : Jurnal Pendidikan Dan Matematika*, *10*(2), 221.

https://doi.org/10.30821/axiom.v10i2.10331
Santoso, H. (2025). *Analisa Kesiapan Sekolah dalam Pemanfaatan Teknologi untuk Pembelajaran*. *14*(1), 223–232.
Sari, I., & Hamami, T. (2022). Pengembangan Metode Flipped Classroom dalam Pendidikan Agama Islam: Solusi Pembelajaran di Masa Pandemi Covid-19. *Edukatif : Jurnal Ilmu Pendidikan*, *4*(4), 5744–5753. https://doi.org/10.31004/edukatif.v4i4.3420
Sharfina, A., & Soraya, I. (2024). Upaya Membentuk Motivasi Belajar Peserta Didik melalui Model Flipped Classroom dalam Pembelajaran PAI. *Jurnal Pendidikan Islam*, *9*(1), 69–87.
Sofianah, M. D., & Sholihah, M. (2024). Model Pembelajaran Flipped Classroom Sebagai Upaya Peningkatan Keterampilan Berbahasa Siswa Di Madrasah Ibtidaiyah. *MODELING: Jurnal Program Studi PGMI*, *11*(1), 774–790.
Sri Lestari. (2021). Peningkatan Kompetensi Pedagogik Guru Melalui Pelatihan Pemanfaatan Media Sosial dan Penerapan Flipped Classroom dalam Hybrid Learning. *Community Education Engagement Journal*, *3*(1), 39–50. https://doi.org/10.25299/ceej.v3i01.7770
Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal Of Medical Education*, *2*, 53–55.
Wijayanto, Z., Sukestiyarno, Y. L., Wijayanti, K., & Pujiastuti, E. (2022). Pembelajaran Flipped Classroom dalam Meningkatkan Self-Regulated Learning Siswa. *Prosiding Seminar Nasional Pascasarjana*, 1241–1247. http://pps.unnes.ac.id/pps2/prodi/prosiding-pascasarjana-unnes
Wright, Benjamin, D, L., & m, J. (1994). Reasonable Mean-Square Fit Values. *Rasch Measurement Transaction*, *8*(3), 370.
Yasin, M., & Judijanto, L. (2024). *Model Pembelajaran Berbasis Teknologi*. PT.Green Pustaka Indonesia. https://books.google.co.id/books?id=LuEPEQAAQBAJ&lpg=PR2&dq=kelas terbalik&lr&pg=PR2#v=onepage&q=kelas terbalik&f=false
Yen, W. M. (1993). Scaling Performance Assessments: Strategies for Managing Local Item Dependence. *Journal Of Education Measurement*, *30*(3), 187–213.
Zwick, R. (2012). *A Review of ETS Differential Item Functioning Assessment Procedures: Flagging Rules, Minimum Sample Size Requirements, and Criterion Refinement*. http://www.ets.org/research/contact.html