Islamic Journal of Integrated Science Education (IJISE), Vol. 5 No. 2, July 2026, pp. 129-156
DOI: https://doi.org/10.30762/ijise.v5i2.8861

Islamic Journal of Integrated Science Education (IJISE)

Program Studi Tadris IPA
Universitas Islam Negeri (UIN) Syekh Wasil Kediri
e-ISSN : 2986-0865
https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/ijise

RASCH-BASED DIAGNOSTIC ANALYSIS OF BIOLOGY EDUCATION STUDENTS' CONCEPTUAL DIFFICULTIES IN WORK AND ENERGY

I Made Hermanto[1]*, Wahyu Mu'zizat Mohamad[2], Lukman Samatowa[3], Ganesha Antarnusa[4]
[1,2,3]Program Studi Pendidikan Fisika, Jurusan Fisika, FMIPA, Universitas Negeri Gorontalo, Indonesia
[4]Gifu University, Japan

***Correspondence E-mail:** imadehermanto@ung.ac.id

Abstract: This study aimed to analyze Biology education students' conceptual difficulties in work and energy using the Rasch Model. A descriptive quantitative approach with a diagnostic survey design was employed. The participants were 21 Biology education students who had learned the topic of work and energy. The instrument consisted of 15 diagnostic multiple-choice items scored dichotomously, with 1 indicating a correct answer and 0 indicating an incorrect answer. The data were analyzed using the Rasch Model to examine instrument quality, item fit, item difficulty, Wright Map, score-category monotonicity, and students' conceptual difficulty patterns. The results showed that students' average ability was higher than the average item difficulty, with a mean person measure of 1.61 logits. However, the person reliability of 0.08 and person separation index of 0.30 indicated that the instrument had limited ability to distinguish students across different ability levels. The most difficult items were Q14, Q3, Q11, Q13, and Q15, which were related to the application of work, potential energy, and mechanical energy conservation. In contrast, Q2, Q4, Q5, and Q9 were the easiest items because they mainly assessed basic conceptual recognition. These findings indicate that students' main difficulties were not in recognizing basic terms but in applying and integrating work and energy concepts in physics contexts.

Keywords: Rasch Model; Conceptual Difficulties; Biology Education Students; Work and Energy

Article History:
Received: 05 May 2026; Revised: 25 July 2026; Accepted: 30 July 2026; Published: 31 July 2026

Citation (APA Style):
Hermanto, I. M., Mohamad, W. M., Samatowa, L., & Antarnusa, G. (2026). Rasch-Based Diagnostic Analysis of Biology Education Students' Conceptual Difficulties In Work and Energy. *Islamic Journal of Integrated Science Education (IJISE), 5*(2), 129–156. https://doi.org/10.30762/ijise.v5i2.8861

INTRODUCTION

Physics is one of the fundamental sciences that Biology education students at universities are required to study in order to develop a scientific understanding of various life phenomena that cannot be fully explained from a biological perspective alone (Adelia et al., 2024; Kaczmarek & Greczyło, 2025). These phenomena are closely related to physics principles such as force, motion, energy, pressure, fluids, waves, electricity, and thermodynamics. In the context of modern science education, the connection between physics and biology has become increasingly important because students are expected to use physical models and quantitative reasoning to analyze biological phenomena in a more integrated and in-depth manner (Munifah et al., 2019; Lestari et al., 2021; Saregar et al., 2025). Physics learning that is designed to be relevant to real-life and biological phenomena can help Biology students develop more complex models, connect multiple scientific ideas, and apply physics equations in new contexts (Crouch & Heller, 2014; Hoskinson et al., 2014; Kaczmarek & Greczyło, 2025; Redish et al., 2014).

One physics topic that is particularly important for Biology students is work and energy. Energy is not only a central concept in physics but also a cross-disciplinary concept that appears in chemistry, biology, ecology, physiology, and environmental science (Dyachenko et al., 2024; Lancor, 2014; Opitz et al., 2019). In biology, energy is related to metabolism, cellular respiration, photosynthesis, muscle movement, membrane transport, food chains, and ecosystem dynamics. However, although the term energy is frequently used in biology, students' understanding of energy from a physics perspective is often incomplete. Students may understand energy as "something used by living organisms," but they may not necessarily be able to explain energy as a physical quantity related to work, changes in kinetic energy, changes in potential energy, mechanical energy conservation, and the work done by conservative and nonconservative forces (Chabalengula et al., 2012; Opitz et al., 2017).

These difficulties become more complex because the concepts of work and energy in physics are abstract, mathematical, and representational. Students are not only required to memorize formulas but also to understand the relationships among force, displacement, angle, energy change, system, environment, and the principle of energy conservation. In several cases, students still experience difficulties in distinguishing work in everyday language from work in physics, determining positive and negative work, understanding work done by friction, explaining the relationship between network and changes in kinetic energy, and applying the law of mechanical energy conservation in situations involving nonconservative forces (Jewett Jr., 2008; Permatasari et al., 2018; Rivaldo et al., 2020). Students often perceive physics as a collection of formulas to be solved mathematically rather than as a conceptual tool for

DOI: https://doi.org/10.30762/ijise.v5i2.8861

explaining biological phenomena. Weak understanding of work and energy may affect students' ability to explain biological processes that involve mechanical and quantitative reasoning, such as energy transformation in body systems, muscle work, energy flow in ecosystems, and the relationship between energy, force, and organism movement. Therefore, physics learning for Biology students should not only emphasize numerical problem solving but also include conceptual diagnosis that can reveal students' specific difficulties and misconceptions.

Several studies in physics education have shown that conceptual understanding is an important indicator of successful physics learning (Laili, et al., 2025; Banda & Nzabahimana, 2021; Fatonah, 2024; Hau & Nuri, 2019). In the topic of work and energy, the need for diagnostic instruments has become increasingly important. Samsudin et al. (2021) developed the Multitier Open-ended Work and Energy Instrument (MOWEI) using Rasch analysis to identify students' misconceptions about work and energy. The instrument covers concepts such as positive and negative work, the law of mechanical energy conservation, the work-energy theorem, and work done by conservative and nonconservative forces. The results showed that MOWEI was valid and reliable for identifying misconceptions, with the highest percentage of misconception reaching 73% on one item. Faihah et al. (2025a) also adapted and validated a four-tier diagnostic instrument on work, power, and energy using the Rasch Model. Their findings indicated that the instrument was reliable, met the assumption of unidimensionality, and showed no item bias based on gender. Samsudin et al. (2025) and Aripiani et al. (2023) also developed tiered diagnostic tests to measure levels of conceptual understanding using Rasch analysis, showing that many students still experienced misconceptions and lacked conceptual understanding of work and energy.

Although many studies have diagnosed students' conceptual understanding of work and energy, most previous research has focused on secondary school students, prospective physics teachers, or physics education students (Firmansyah et al, 2023; Fatonah & Wicaksana, 2023; Nur'aini et al, 2023). Studies that specifically analyze Biology education students' conceptual difficulties in work and energy remain relatively limited, particularly those using modern measurement approaches such as the Rasch Model. Biology students have different characteristics from physics students. They are more familiar with energy in the context of metabolism, organ function, organisms, and ecosystems, but they may not necessarily have a strong understanding of energy as a physical quantity that can be mathematically modeled. This gap makes research on Biology education students important, as it can provide a more specific picture of the conceptual barriers that hinder their understanding of basic physics.

The Rasch Model is also highly appropriate for this study because it provides stronger diagnostic information than raw score analysis or classical test theory. The Rasch Model enables researchers to map students' ability and item difficulty on the same logit scale, examine item fit with the measured construct, identify items that are too easy or too difficult, assess person and item reliability, and detect potential item bias through Differential Item Functioning or DIF (Medvedev Oleg N. & Krägeloh, 2025; Sumintono & Widhiarso, 2015). In physics education, the Rasch Model is considered important because it produces linear measures of person ability and item difficulty and helps evaluate instrument quality as well as inconsistent response patterns (Lamanepa, 2021; Taasoobshirazi et al., 2015). National studies have also shown that the Rasch Model is widely used for validating science instruments, including mapping item difficulty levels, evaluating reliability, analyzing item fit, and examining bias based on respondent characteristics (Aripiani et al., 2023; Faihah et al., 2025a; Samsudin et al., 2021, 2025).

Based on this background, this study aims to determine the level of students' conceptual mastery and to map the hierarchy of conceptual difficulties in the subtopics of work and energy. Through Rasch analysis, this study reveals which items are the most difficult, which sub-concepts are most prone to misconceptions, how students' ability is distributed relative to item difficulty, and whether the instrument used has adequate psychometric quality. The findings are expected to provide theoretical contributions to the development of diagnostic assessment in interdisciplinary physics education, as well as practical contributions for basic physics lecturers in designing instruction that is more adaptive, contextual, and relevant for Biology students.

METHOD

Research Design

This study employed a descriptive quantitative approach with a diagnostic survey design. This design was selected because the study aimed to analyze Biology Education students' conceptual difficulties in understanding work and energy, while also mapping the relationship between students' ability and item difficulty using the Rasch Model. The Rasch approach is relevant in science education research because it places person ability and item difficulty on the same logit scale. Therefore, the analysis does not merely indicate whether students' answers are correct or incorrect, but also provides diagnostic information regarding item quality, the distribution of students' ability, and the hierarchy of conceptual difficulty.

Participants

The study involved 21 Biology Education students who had previously studied work and energy in a Basic Physics course. The participants consisted of 19 female students (90.48%) and 2 male students (9.52%). The participants in this study were students from a Biology or Biology Education study program who had previously received instruction on work and energy in a Basic Physics course. The participants were selected based on the criteria that they had studied the topic, were present during data collection, and completed all test items. The Biology Education context was chosen because students in this program study physics as a foundational science that supports biological understanding. Therefore, it is important to examine the extent to which they understand physics concepts related to life phenomena, such as energy, motion, muscle work, and energy transformation in biological systems.

Research Instrument

The instrument consisted of 15 diagnostic multiple-choice items distributed across seven content domains. Three items assessed basic concepts of work, the unit of work, and zero work (Q1, Q2, and Q10); three items assessed the calculation of work involving force, displacement, and force direction (Q3, Q11, and Q14); one item assessed work from a force-position graph (Q12); one item assessed the basic concept of energy (Q4); three items assessed gravitational potential energy (Q5, Q8, and Q13); two items assessed kinetic energy (Q6 and Q7); and two items assessed mechanical energy and energy conservation (Q9 and Q15).

Table 1. Distribution of test items across content domains

Content domain	Item numbers	Number of items
Basic work concepts, unit, and zero work	Q1, Q2, Q10	3
Work calculation involving force and displacement	Q3, Q11, Q14	3
Work from a force-position graph	Q12	1
Basic concept of energy	Q4	1
Gravitational potential energy	Q5, Q8, Q13	3
Kinetic energy	Q6, Q7	2
Mechanical energy and energy conservation	Q9, Q15	2
Total		15

Each item contained five answer options consisting of one correct answer and four distractors. The construction of distractors was an important component of the diagnostic test because well-designed distractors can help reveal students' conceptual difficulties and alternative conceptions. The instrument was reviewed by one expert in physics education before it was administered to the participants. The expert review examined the alignment between each item and its intended conceptual indicator, the accuracy of the physics content, the clarity of the wording, the consistency of symbols and units, the suitability of the figures, and the plausibility of the answer options and distractors. The qualitative comments provided by the expert were used to revise the items before data collection.

The main revisions involved clarifying items that referred to figures or reference points, correcting ambiguous or overly general definitions, standardizing physical quantities and SI units, and ensuring that numerical information was sufficient for solving the problem. For example, the reference level in the gravitational potential energy item was clarified, while items involving force direction were revised to make the angle between force and displacement explicit. Because the validation was conducted as an expert content review, no statistical content-validity coefficient is reported.

Data Collection Procedure

Data were collected after students had completed instruction on the topic of work and energy. Students completed the 15-item diagnostic test independently within 90 minutes under the supervision of the research team. Before administering the test, the researcher explained the general purpose of the study, the test instructions, and the assurance that students' responses would be used solely for academic purposes. Students' responses were scored dichotomously, with a score of 1 assigned to a correct answer and a score of 0 assigned to an incorrect answer. The resulting person-item response matrix was analyzed using Ministep software.

Data Analysis

Rasch analysis was conducted after the completeness of the response data had been checked and all responses had been coded dichotomously. The person-item matrix was then imported into Ministep. The analysis proceeded through the following stages, (1) examination of person and item summary statistics, reliability, and separation; (2) assessment of unidimensionality and residual local dependence; (3) examination of the functioning and monotonicity of the dichotomous score categories; (4) item-fit analysis using infit MNSQ, outfit MNSQ, and point-measure correlation; (5) estimation of item difficulty in logits; (6) construction of the Wright Map; (7) examination of person-fit statistics; and (8) interpretation of the results by connecting item measures with the conceptual content represented by each item. The interpretation of the results was carried out by linking item logit values, students' response patterns, and the conceptual indicators represented by each item. Items with high difficulty levels were interpreted as representing concepts that were most difficult for students to understand. For example, if items related to gravitational potential energy or mechanical energy conservation had high logit values, these sub-concepts were categorized as major areas of conceptual difficulty. Thus, the analysis did not only describe students' overall conceptual mastery, but also provided diagnostic information about specific parts of work and energy that require further instructional emphasis in Basic Physics learning for Biology Education students.

Ethical Considerations

This study was conducted in accordance with ethical principles in educational research. Students' identities were kept confidential, participation was voluntary, and the test results were not used to affect students' academic grades. The data obtained were used only for scientific analysis and instructional development purposes.

FINDING AND DISCUSSION

This section presents the results and discussion of Biology Education students' conceptual difficulties in work and energy based on a diagnostic multiple-choice test and Rasch Model analysis. The results are organized into several sections, including respondents' characteristics, Rasch summary statistics, unidimensionality and local dependence, score-category monotonicity, item fit, item difficulty hierarchy, Wright Map, person fit, and a synthesis of students' conceptual difficulties. This structure is intended to provide a comprehensive picture of both the psychometric quality of the instrument and the conceptual areas in which Biology Education students experienced difficulties in understanding work and energy.

Respondents' Characteristics and Response Data Completeness

This study involved 21 Biology Education students who completed 15 diagnostic multiple-choice items on work and energy. All student responses were scored dichotomously, with a score of 1 assigned to a correct answer and a score of 0 assigned to an incorrect answer. Based on gender characteristics, the sample was dominated by female students, consisting of 19 students or 90.48%, whereas male students accounted for 2 students or 9.52%. All 21 participants completed all 15 items, producing 315 valid responses with no missing data. This information is reported to document the completeness of the person-item matrix before Rasch estimation.

Table 2. Summary of response data completeness

Data Component	**Value**
Number of respondents	21 students
Number of items	15 items
Expected total responses	315 responses
Missing responses	0
Valid responses for analysis	315 responses
Data completeness percentage	100%

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Considering the sample size, this study is more appropriately positioned as an initial diagnostic study of Biology Education students' conceptual difficulties rather than as a large-scale final calibration of the instrument. Nevertheless, this does not reduce the value of the study, as Rasch studies with limited samples can still provide preliminary information about students' ability patterns, item difficulty levels, and items requiring revision. In science education, the Rasch Model is widely used not only to generate scores but also to examine instrument quality and conceptual difficulty patterns in a more structured manner (Planinic et al., 2019).

Rasch Summary Statistics

The Rasch summary statistics indicate that students' average ability was higher than the average item difficulty. The mean person measure was 1.61 logits, whereas the mean item difficulty was set at 0.00 logits. This suggests that, in general, the test items were relatively easy for the group of students examined. However, the person reliability value of 0.08 and the person separation index of 0.30 indicate that the instrument was not sufficiently strong in distinguishing students into clearly separated ability strata. From the item perspective, the item reliability value of 0.62 and item separation index of 1.28 suggest that the data were beginning to differentiate easier items from more difficult ones, although the stability of the item difficulty hierarchy was not yet strong.

The low person reliability and person separation index may be explained by several factors. First, the limited number of respondents may have restricted the range of ability captured by the model. Second, the relatively small number of items, namely 15, may have limited the precision of person ability estimates. Third, most items appeared to be below the average ability level of the students, making the instrument less optimal for distinguishing students with moderate and high levels of conceptual ability. In the Rasch Model, an instrument that is too easy for a given group of respondents tends to provide limited information for separating ability levels.

Table 3. Rasch summary statistics for persons and items

Rasch Parameter	Person	Item
Mean measure	1.61 logits	0.00 logits
Standard deviation	0.80	1.23
Reliability	0.08	0.62
Separation index	0.30	1.28
Mean infit MNSQ	1.01	1.00
Mean outfit MNSQ	0.89	0.89

Nevertheless, this condition does not mean that the instrument is unusable. The mean infit MNSQ values for persons and items were 1.01 and 1.00, respectively, indicating that, in general, students' response patterns and item characteristics were consistent with the expectations of the Rasch Model.

Table 4. Operational criteria used for rasch item interpretation

Parameter	Operational criterion	Interpretation
Infit MNSQ	0.50–1.50	Values within this range indicate that the item is productive for measurement. Values below 0.50 indicate overfit or overly predictable responses, whereas values above 1.50 indicate underfit or excessive unexpected responses.
Outfit MNSQ	0.50–1.50	Values within this range indicate acceptable sensitivity to unexpected or outlying responses. Values outside the range require further statistical and substantive review.

(Laliyo et al, 2022; Schutte et al, 2016)

The mean outfit MNSQ value of 0.89 was also within an acceptable range. In Rasch measurement literature, fit statistics close to 1.00 indicate that response variations are reasonably consistent with model expectations, whereas extreme deviations require further examination at the item or person level (Planinic et al., 2019).

These findings are consistent with Rasch studies in physics and science education, which show that an instrument can still provide meaningful diagnostic information even when person separation is not high, particularly when the purpose of the study is to identify item difficulty hierarchies and problematic conceptual patterns. Studies on students' conceptual understanding of static fluids using Rasch analysis also suggest that participants' ability distribution may be higher than the average item difficulty; therefore, the analysis should not focus solely on scores but also on the alignment between participant ability and item difficulty (Khusaini et al., 2025).

Thus, the instrument provides useful preliminary information regarding students' conceptual difficulties, although further refinement is required if it is to be used for more precise measurement. Adding items with moderate to high difficulty levels would improve the measurement coverage, particularly in distinguishing students who merely recognize basic concepts from those who can apply work and energy concepts in more complex contexts (Nurulsari et al., 2020; Usman et al., 2025; Putra et al., 2024).

Instrument Unidimensionality and Local Dependence

The unidimensionality analysis using Principal Component Analysis of Rasch Residuals showed that the raw variance explained by measures was 24.2%, close to the expected value of

25.2%. The variance explained by persons was 8.4%, whereas the variance explained by items was 15.8%. These results provide preliminary evidence that the instrument measured the same main construct, namely Biology Education students' conceptual understanding of work and energy. However, the strength of unidimensionality cannot yet be considered high.

Particular attention should be given to the unexplained variance in the first contrast, which reached 3.7919 eigenvalue. This value indicates the possibility of an additional residual pattern beyond the main construct. The contrast loading results also showed that items Q2, Q5, and Q9 had very high positive loadings of 0.94, whereas Q13 and Q15 were located on the negative side. This pattern suggests that some items may form a conceptual cluster or share similar wording, context, or solution strategies. This finding is further supported by very high residual correlations between item pairs Q2–Q5 and Q2–Q9, each reaching 1.00, indicating strong local item dependence. Local item dependence was examined using standardized residual correlations generated by Ministep. Positive residual correlations approaching or exceeding 0.70 were interpreted as evidence of substantial local dependence. Therefore, the residual correlations of 1.00 between Q2-Q5 and Q2-Q9 were not used as a cut-off criterion; rather, they were observed values that clearly exceeded the operational threshold and indicated nearly identical residual response patterns. This result suggests that these items may provide redundant information because they predominantly assess direct recognition of basic concepts.

In the Rasch Model, unidimensionality and local independence are important assumptions because each item should ideally provide relatively independent information about the measured construct. When several items remain highly correlated after controlling for general ability, they may provide redundant information and weaken the interpretation of measurement results. Planinic et al. (2019) emphasized that Rasch analysis is not only used to estimate person ability and item difficulty but also to evaluate whether the data support the expected measurement structure.

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Table 5. Results of instrument unidimensionality analysis

Analysis Component	Eigenvalue	Observed	Expected	Interpretation
Total raw variance in observations	19.7952	100.0%	100.0%	Total variance of the data
Raw variance explained by measures	4.7952	24.2%	25.2%	Variance explained by the Rasch model
Raw variance explained by persons	1.6664	8.4%	8.8%	Contribution of students' ability
Raw variance explained by items	3.1288	15.8%	16.5%	Contribution of item difficulty
Raw unexplained variance total	15.0000	75.8%	74.8%	Residual variance not explained by the model
Unexplained variance in the 1st contrast	3.7919	19.2%	25.3%	Possible first residual dimension
Unexplained variance in the 2nd contrast	2.5266	12.8%	16.8%	Additional residual indication
Unexplained variance in the 3rd contrast	2.2600	11.4%	15.1%	Additional residual indication
Unexplained variance in the 4th contrast	1.6575	8.4%	11.1%	Relatively lower
Unexplained variance in the 5th contrast	1.2436	6.3%	8.3%	Relatively low
Essential unidimensionality	—	31.9%	—	Limited unidimensionality

This finding is also consistent with Hofer et al. (2017), who used Rasch analysis as part of the validation process in developing a conceptual mechanics instrument to ensure that test items functioned consistently in measuring conceptual understanding. In the Indonesian context, the use of Rasch analysis in science diagnostic instrument development and validation has also shown that examining item quality, reliability, and response structure is essential to ensure that an instrument measures the intended construct (Sekarningtias et al., 2023; Hardini et al., 2026).

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Based on these results, the instrument can still be used for diagnostic analysis, but interpretation should be made cautiously. Items Q2, Q5, and Q9 should be reviewed to ensure that they do not merely measure the same basic concept repeatedly. If these items are related to definitions or recognition of basic terms, revision may be directed toward differentiating indicators, varying contexts, or increasing the reasoning demand. Meanwhile, Q13 and Q15, which appeared on the opposite side of the residual structure, may represent more complex conceptual aspects, such as potential energy, energy transformation, or mechanical energy conservation.

Monotonicity of the Test Items

The monotonicity analysis showed that the dichotomous score categories in the instrument functioned according to the basic assumptions of the Rasch Model. Category 0, representing an incorrect response, comprised 69 responses or 22% of the data and had an observed average measure of 0.38 logits. Category 1, representing a correct response, comprised 246 responses or 78% and had an observed average measure of 2.01 logits.

Table 6. Summary of the dichotomous category structure

Score category	Interpretation	Frequency	Percentage	Observed average measure	Infit MNSQ	Outfit MNSQ
0	Incorrect response	69	22%	0.38	1.01	0.85
1	Correct response	246	78%	2.01	1.01	1.01

The increase in the observed average from category 0 to category 1 indicates that students who answered correctly generally had higher estimated ability than those who answered incorrectly. The infit and outfit MNSQ values were 1.01 and 0.85 for category 0 and 1.01 and 1.01 for category 1, respectively, indicating that both categories functioned productively.

The dichotomous curve also showed a pattern supporting monotonicity. When students' ability was lower than item difficulty, the probability of obtaining a score of 0 was higher. Conversely, when students' ability increased and exceeded item difficulty, the probability of obtaining a score of 1 increased. The intersection point of the two categories was located around zero on the person-minus-item measure scale, indicating that the probability of answering correctly and incorrectly was relatively balanced when students' ability was equivalent to item difficulty.

```
DICHOTOMOUS CURVES
P      -+--------------+--------------+--------------+--------------+-
R  1.0 +                                                             +
O      |                                                             |
B      |0                                                           1|
A      | 000000                                               111111 |
B   .8 +       00000                                     11111       +
I      |            0000                             1111            |
L      |                0000                     1111                |
I      |                    000               111                    |
T   .6 +                       000         111                       +
Y      |                          000   111                          |
    .5 +                             ***                             +
O      |                          111   000                          |
F   .4 +                       111         000                       +
       |                    111               000                    |
R      |                1111                     0000                |
E      |            1111                             0000            |
S   .2 +       11111                                     00000       +
P      | 111111                                               000000 |
O      |1                                                           0|
N      |                                                             |
S   .0 +                                                             +
E      -+--------------+--------------+--------------+--------------+-
        -2             -1              0              1              2
   PERSON [MINUS] ITEM MEASURE
```

Figure 1. Monotonicity Curve of Dichotomous Score Categories in the Rasch Model

This pattern is consistent with the basic principle of the Rasch Model, in which the probability of a correct response is determined by the difference between person ability and item difficulty. In physics education research, this principle is important because it enables researchers to interpret students' responses not simply as right or wrong but as probabilistic indicators of the position of students' conceptual ability relative to item difficulty (Planinic et al., 2019).

Although the score categories functioned monotonically, the proportion of correct answers reached 78%, indicating that the instrument tended to be relatively easy for the respondents. This issue should be considered because incorrect answers in a multiple-choice diagnostic test do not always indicate low general ability. In a diagnostic context, incorrect answers may represent misconceptions, difficulties in reading representations, reliance on everyday intuition, or the influence of distractors. Mulyani et al. (2021) showed that Rasch-based distractor analysis can help interpret answer choices as indications of misconceptions rather than merely as random errors.

Therefore, the monotonicity results support the adequacy of the dichotomous scoring categories, but the diagnosis of conceptual difficulties should still be connected to item fit, item difficulty, the Wright Map, and distractor analysis. In the topic of work and energy, students may answer definitional items correctly but still experience difficulties when the same concepts appear in quantitative or representational contexts.

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Item Fit Analysis

The item fit analysis showed that all items had infit MNSQ values within the acceptable range, namely between 0.79 and 1.31.

Table 7. Item fit analysis based on the rasch model

Item	Conceptual Category	Measure Logit	Infit MNSQ	Outfit MNSQ	Pt-Measure Corr.	Decision
Q1	Basic concept of work	-0.75	1.13	1.22	0.09	Revise
Q2	Basic concept of work	-1.53	0.93	0.47	0.26	Revise
Q3	Calculation of work	1.48	0.94	0.90	0.48	Fit
Q4	Basic concept of energy	-1.53	0.93	0.47	0.26	Revise
Q5	Gravitational potential energy	-1.53	0.93	0.47	0.26	Revise
Q6	Kinetic energy	-0.75	1.05	1.50	0.12	Revise
Q7	Kinetic energy	-0.25	0.81	0.60	0.43	Fit
Q8	Gravitational potential energy	-0.75	1.09	0.97	0.15	Revise
Q9	Mechanical energy and energy conservation	-1.53	0.93	0.47	0.26	Revise
Q10	Basic concept of work	0.73	1.02	0.92	0.36	Fit
Q11	Calculation of work	1.48	0.98	1.04	0.44	Fit
Q12	Work from a force-position graph	0.73	1.17	1.22	0.22	Fit
Q13	Gravitational potential energy	1.24	1.31	1.29	0.18	Revise
Q14	Calculation of work	1.72	0.99	1.04	0.45	Fit
Q15	Mechanical energy and energy conservation	1.24	0.79	0.73	0.58	Fit

However, several items showed outfit MNSQ values that were too low, namely 0.47, or had low point-measure correlations, such as Q1, Q6, Q8, and Q13. Based on the analysis, seven items were categorized as fit, namely Q3, Q7, Q10, Q11, Q12, Q14, and Q15. Meanwhile, eight items required revision or further review, namely Q1, Q2, Q4, Q5, Q6, Q8, Q9, and Q13.

Items categorized as fit indicate that students' response patterns to those items were relatively consistent with the predictions of the Rasch Model. In other words, students with higher ability tended to have a greater probability of answering these items correctly. Items Q3, Q11, Q14, and Q15 are particularly important because, in addition to meeting the fit criteria, they were also located at higher difficulty levels. Such items have strong diagnostic value because they can reveal concepts that have not yet been mastered by some students.

In contrast, items with low point-measure correlations require closer examination. A low correlation indicates that responses to the item are not sufficiently aligned with the overall ability estimate. This may occur when an item has ambiguous wording, overly strong distractors, unfamiliar contexts, or measures an aspect outside the intended construct. For items Q2, Q4, Q5, and Q9, the outfit MNSQ value of 0.47 indicates a response pattern that was overly predictable. A low outfit value does not necessarily mean that the item is poor, but it may indicate that the item is too easy or contributes limited additional information to the measurement.

This finding is consistent with Rasch literature, which views fit statistics as tools for identifying items that require review rather than as a sole basis for accepting or rejecting items. Planinic et al. (2019) emphasized that item fit in physics education research helps researchers understand whether items function according to the intended construct and whether participants' response patterns are consistent with the expected hierarchy of item difficulty. Hofer et al. (2017) also showed that the development of conceptual mechanics tests requires a combination of quantitative Rasch evaluation and qualitative content review to ensure that the instrument truly represents conceptual understanding.

In the national context, Rasch analysis has also been used to evaluate item quality in diagnostic instruments for misconceptions and science conceptual understanding. A study on Rasch-based analysis of chemistry misconceptions showed that items and distractors should be interpreted together so that items that appear "easy" or "fit" statistically can still be evaluated for their diagnostic value (Mulyani et al., 2021). Therefore, the items marked for revision in this study should not necessarily be removed immediately, but should be re-examined in terms of conceptual indicators, wording, context, and the functioning of answer options.

Item Difficulty and Conceptual Difficulty

The item difficulty analysis showed that logit measures ranged from -1.53 to +1.72 logits. The most difficult item was Q14, with a measure of +1.72 logits, followed by Q3 and Q11, each

DOI: https://doi.org/10.30762/ijise.v5i2.8861

with a measure of +1.48 logits, and Q13 and Q15, each with a measure of +1.24 logits. Conversely, the easiest items were Q2, Q4, Q5, and Q9, each with a measure of -1.53 logits.

Table 8. Hierarchy of item difficulty based on logit measures

Difficulty Rank	Item	Measure Logit	Interpretation
1	Q14	+1.72	Very difficult
2	Q3	+1.48	Very difficult
3	Q11	+1.48	Very difficult
4	Q13	+1.24	Very difficult
5	Q15	+1.24	Very difficult
6	Q10	+0.73	Difficult
7	Q12	+0.73	Difficult
8	Q7	-0.25	Moderate
9	Q1	-0.75	Easy
10	Q6	-0.75	Easy
11	Q8	-0.75	Easy
12	Q2	-1.53	Very easy
13	Q4	-1.53	Very easy
14	Q5	-1.53	Very easy
15	Q9	-1.53	Very easy

This hierarchy indicates that students' difficulties were not primarily located in recognizing basic terms, but rather in applying and integrating concepts. Items Q2, Q4, Q5, and I9 were likely related to basic concepts such as the unit of work, the definition of energy, potential energy, and mechanical energy as the combination of kinetic and potential energy. These items were relatively easy because they could be answered using declarative knowledge. In contrast, items Q14, Q3, and Q11, which were related to the calculation of work, required students to connect force, displacement, and the possible direction of force. Items Q13 and Q15, which were related to potential energy and mechanical energy conservation, required students to understand relationships among quantities within an energy system.

This finding is consistent with Tong et al. (2023), who showed that work and mechanical energy require complex knowledge integration. Students' difficulties are not only related to formula mastery, but also to their ability to connect work, changes in kinetic energy, mechanical energy, and the system context being analyzed. Similar findings were also reported in the development of MOWEI and FORTUNE instruments, which showed that work and energy are topics prone to misconceptions and therefore require diagnostic instruments capable of distinguishing conceptual understanding, confidence in answers, and alternative conceptual patterns (Aripiani et al., 2023; Samsudin et al., 2021).

In the context of Biology Education students, difficulties with work calculation and mechanical energy conservation may arise because students are more familiar with energy in biological contexts, such as metabolism or energy flow, than with mechanical relationships among force, displacement, and energy change. Therefore, these findings suggest that learning work and energy for non-physics students should more explicitly connect mathematical representations with conceptual meanings and phenomena that are familiar to their disciplinary background.

Wright Map of Students' Ability and Item Difficulty

The Wright Map showed that students' ability ranged from approximately 0 to 4 logits. The students with the highest ability were P1 and P2, both located around 4 logits, whereas the student with the lowest ability was P21, located close to 0 logits. From the item perspective, Q14 was positioned at the highest difficulty level, followed by Q11, Q3, Q13, and Q15. In contrast, Q2, Q4, Q5, and Q9 were positioned at the lowest difficulty level, around -1.53 logits.

This map shows that most students had ability levels above the average item difficulty. This condition indicates that the instrument was relatively easy for most respondents. Nevertheless, the presence of Q14, Q11, Q3, Q13, and Q15 at higher difficulty levels still provides important information regarding the most challenging sub-concepts. Thus, the Wright Map reveals two important findings simultaneously: students were generally able to answer many items correctly, but there remained a cluster of concepts requiring higher-order reasoning.

This finding is consistent with Rasch-based research on conceptual understanding of static fluids, which showed that participants' ability distribution may be above the average item difficulty, while the person-item map remains useful for identifying the most challenging items and ability regions not sufficiently covered by the instrument (Khusaini et al., 2025). In physics education research, the Wright Map is considered important because it provides a visual representation of the alignment between participants' ability and item difficulty. Planinic et al. (2019) also emphasized that mapping persons and items on the same logit scale is one of the key advantages of Rasch analysis over raw-score-based approaches.

Islamic Journal of Integrated Science Education (IJISE), Vol. 5 No. 2, July 2026, pp. 129-156
DOI: https://doi.org/10.30762/ijise.v5i2.8861

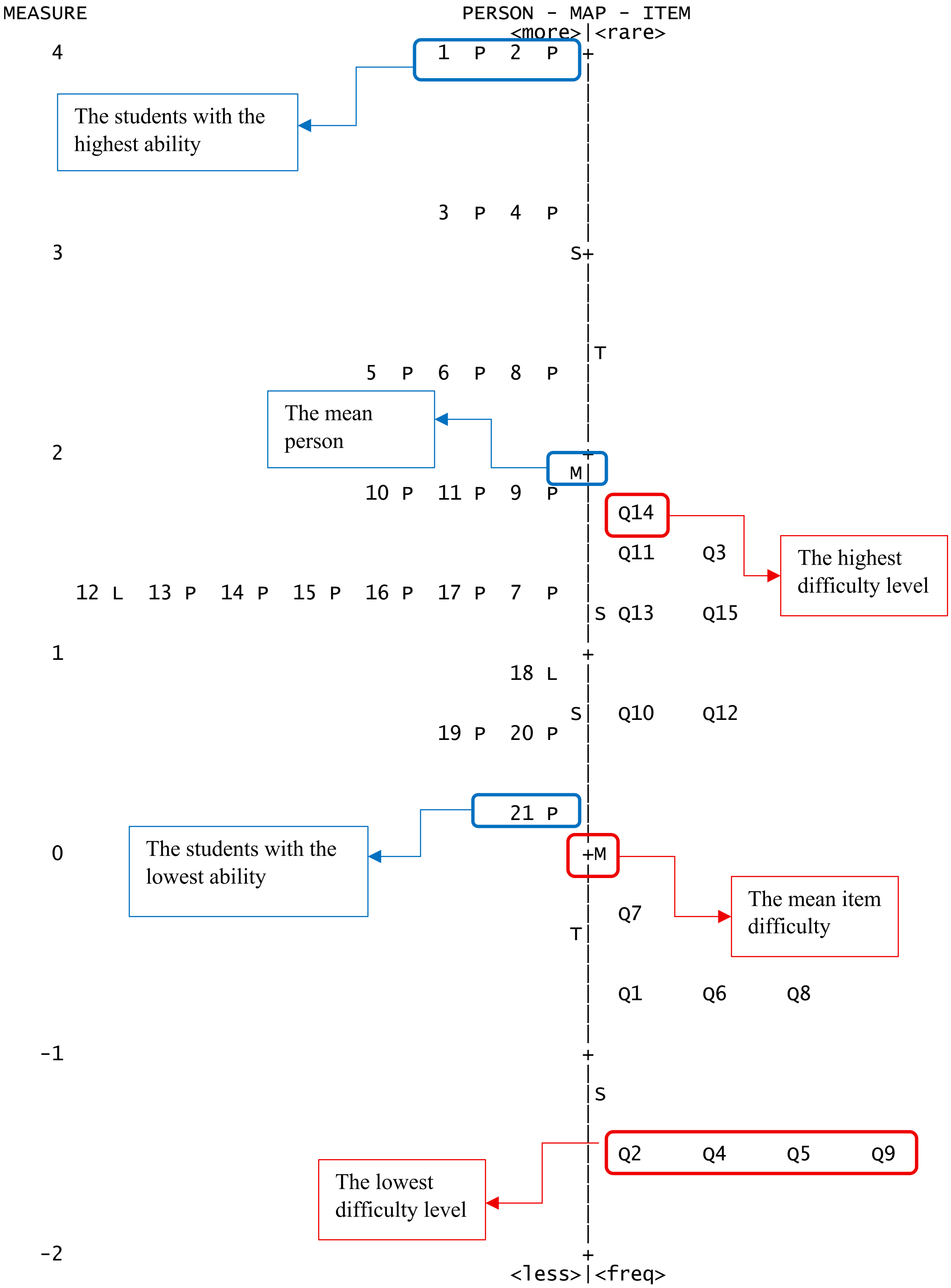

Figure 2. Wright Map of Students' Ability and Item Difficulty in Work and Energy

The Wright Map in this study indicates an imbalance between the distribution of students' ability and the distribution of item difficulty. Many students were located above the average item difficulty, while items with moderate to high difficulty levels were relatively limited. In diagnostic instrument development, this condition suggests the need for additional items in the range of approximately 0 to +2 logits so that the instrument can better capture variations in students' ability at moderate and high levels. Additional items may be designed to require students to interpret diagrams, compare energy in two states, or explain the relationship between work and energy change in biological contexts.

Person Fit and Consistency of Students' Response Patterns

The person fit analysis showed that most students had response patterns that were acceptable within the Rasch Model. However, several respondents required closer attention. Student P20 showed the most misfitting response pattern, with an infit MNSQ of 1.63, outfit MNSQ of 2.20, and point-measure correlation of 0.01. Student P8 also showed an indication of misfit, with an outfit MNSQ of 2.02 and point-measure correlation of 0.00. Meanwhile, P19 was categorized as marginally misfitting, with an infit MNSQ of 1.52 and outfit MNSQ of 1.63.

High outfit values indicate unexpected responses, such as a student answering relatively easy items incorrectly while answering more difficult items correctly. In the context of multiple-choice tests, such patterns may be caused by guessing, misreading the item, lack of attention, or unstable conceptual understanding. P8 is particularly noteworthy because this student obtained a relatively high total score of 13 out of 15, yet the response pattern was not fully consistent with the model prediction. This indicates that a high score does not always correspond to a stable pattern of conceptual understanding.

Conversely, several students, such as P3, P4, P10, P12, and P18, showed overfit patterns, meaning that their responses were overly predictable. Overfit is usually less problematic than high misfit, but it still suggests that students' responses strongly followed the item difficulty order and therefore provided limited additional diagnostic information. In Rasch research, person fit analysis is useful for identifying unusual response patterns and helping researchers decide whether certain data require further examination (Planinic et al., 2019).

This finding is consistent with Rasch-based diagnostic studies emphasizing that participants' response patterns should be interpreted as part of data validity rather than merely as individual scores. Research on misconception diagnostic instruments suggests that inconsistent responses may provide evidence of fragmented reasoning, low confidence in answers, or guessing strategies (Mulyani et al., 2021). Therefore, students such as P20, P8, and P19 should not necessarily be removed from the analysis, but they may be considered cases that reflect variations in conceptual understanding patterns.

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Table 9. Person fit analysis

Person	Total Score	Measure Logit	Infit MNSQ	Outfit MNSQ	Pt-Measure Corr.	Decision
P20	9	0.55	1.63	2.20	0.01	Misfit
P8	13	2.37	1.34	2.02	0.00	Misfit
P19	9	0.55	1.52	1.63	0.16	Marginal misfit
P9	12	1.81	1.31	1.42	0.17	Fit, but requires attention
P14	11	1.35	1.19	1.11	0.33	Fit
P5	13	2.37	1.08	0.70	0.31	Fit
P16	11	1.35	1.08	1.05	0.39	Fit
P7	11	1.35	1.03	1.01	0.42	Fit
P11	12	1.81	0.96	0.61	0.47	Fit
P3	14	3.22	0.94	0.44	0.32	Overfit
P15	11	1.35	0.93	0.76	0.51	Fit
P17	11	1.35	0.93	0.63	0.54	Fit
P6	13	2.37	0.91	0.50	0.43	Fit
P13	11	1.35	0.87	0.59	0.58	Fit
P4	14	3.22	0.86	0.36	0.37	Overfit
P10	12	1.81	0.68	0.41	0.64	Overfit
P21	8	0.18	0.64	0.57	0.78	Fit
P12	11	1.35	0.61	0.42	0.73	Overfit
P18	10	0.94	0.61	0.46	0.77	Overfit

In the article, person fit results should be reported as evidence that most response data were sufficiently supportive of Rasch analysis, while several response patterns still required cautious interpretation. In future studies, respondents with high misfit values may be selected for diagnostic interviews to investigate whether their errors resulted from misconceptions, guessing, representational difficulties, or non-cognitive factors such as carelessness during test-taking.

Synthesis of Students' Conceptual Difficulties

Overall, the findings indicate that Biology Education students' conceptual difficulties in work and energy were more prominent in the application and integration of concepts than in the recognition of basic definitions. Very easy items, such as Q2, Q4, Q5, and Q9, indicate that

DOI: https://doi.org/10.30762/ijise.v5i2.8861

students were relatively able to recognize basic concepts such as the unit of work, the definition of energy, potential energy, and mechanical energy. In contrast, items requiring the application of relationships among force, displacement, work, potential energy, kinetic energy, and mechanical energy conservation, such as Q14, Q3, Q11, Q13, and Q15, were the most difficult items.

Table 10. Item difficulty measures and classification of the developed test items

Item	Abbreviated item content	Conceptual domain	Measure	Classification
Q2	Identifying the SI unit of work	Unit of work	−1.53	Very easy
Q4	Identifying the definition of energy	Basic concept of energy	−1.53	
Q5	Identifying energy associated with an object's position	Gravitational potential energy	−1.53	
Q9	Identifying mechanical energy as the sum of kinetic and potential energy	Mechanical energy	−1.53	
Q3	Calculating work by a 20 N force over 8 m at 60°	Work calculation with force direction	+1.48	Very difficult
Q11	Calculating work by a 30 N horizontal force over 1.5 m	Work calculation	+1.48	
Q13	Determining gravitational potential energy at point B relative to point A	Potential energy and reference level	+1.24	
Q14	Calculating work by a 50 N force at 60° over 4 m	Work calculation with force direction	+1.72	
Q15	Calculating kinetic energy during free fall at points A and B	Mechanical-energy conservation	+1.24	

This pattern reveals a gap between declarative knowledge and integrated conceptual understanding. Students may recognize terms or basic formulas, but they may not necessarily be able to use these concepts to analyze physical situations. Such difficulties are common in the topic of work and energy because the concepts require students to simultaneously understand systems, changes of state, direction of force, displacement, and energy transformation. Tong et al. (2023) showed that students often experience fragmented knowledge when learning work and mechanical energy; they may remember formulas but fail to establish meaningful relationships between work and energy changes.

This finding is also consistent with the development of diagnostic instruments for work and energy, showing that this topic is prone to misconceptions and requires richer diagnostic assessment than ordinary multiple-choice tests (Aripiani et al., 2023; Samsudin et al., 2021). In the Indonesian context, studies adapting four-tier diagnostic tests for work, power, and energy have also emphasized that Rasch-based instruments can be used to map item difficulty and evaluate instrument suitability across cultural contexts (Faihah et al., 2025b).

For Biology Education students, these difficulties can be understood as a consequence of the interdisciplinary nature of the topic. The concept of energy is highly familiar in biology, but it often appears in the form of metabolic energy, energy flow in ecosystems, or energy transformation in organisms. In physics, however, energy in the topic of work and energy is understood through mechanical relationships among force, displacement, system state, and energy change. This difference in context may cause students to recognize the term "energy" in a general sense while experiencing difficulties when applying it in a mechanics-based model.

The findings of this study suggest that instruction on work and energy for Biology Education students should emphasize connections across representations. Formulas should not be taught merely as computational procedures, but should be linked to force diagrams, object position, changes in height, changes in velocity, and the physical meaning of work. The use of biological contexts, such as muscle work, organism movement, lifting activities, or energy transformation in body movement, may help students build a bridge between physics concepts and their disciplinary field. Such an approach can make the learning of work and energy not only mathematical, but also conceptual and meaningful for Biology Education students.

CONCLUSION

This study concludes that Biology Education students generally demonstrated an initial understanding of work and energy, as indicated by the average person ability being higher than the average item difficulty. However, this understanding was not evenly distributed across all sub-concepts. Students found it relatively easier to recognize basic concepts, such as the unit of work, the definition of energy, potential energy, and mechanical energy. In contrast, they experienced greater difficulty with items requiring the application and integration of concepts, particularly those involving the relationships among force, displacement, work, potential energy, kinetic energy, and mechanical energy conservation.

The Rasch analysis revealed that the most difficult items were Q14, Q3, Q11, Q13, and Q15, which represented more complex conceptual demands. These items required students not only to recall formulas but also to connect physical quantities and interpret situations involving work and energy transformation. Conversely, Q2, Q4, Q5, and Q9 were the easiest items because they mainly assessed recognition of basic concepts. These findings indicate that the main conceptual difficulty among Biology Education students lies not in remembering definitions, but in applying and integrating physics concepts within mechanics-based contexts.

From the perspective of instrument quality, the diagnostic test provided preliminary evidence for measuring students' conceptual understanding of work and energy. The dichotomous scoring categories functioned monotonically, and several items showed acceptable fit to the Rasch Model. Nevertheless, the instrument still requires further refinement, as indicated by low person reliability and separation, evidence of local item dependence among several items, and the presence of items requiring revision based on item fit and point-measure correlation. Therefore, the instrument can be used as an initial diagnostic tool, but further development is needed through item revision, the addition of items with moderate to high difficulty levels, and testing with a larger and more diverse sample.

Overall, this study highlights the usefulness of the Rasch Model in diagnosing conceptual difficulties in interdisciplinary physics learning. For Biology Education students, work and energy should not be taught merely as mathematical procedures, but as interconnected concepts that explain biological and physical phenomena. Future instruction should emphasize multiple representations, conceptual reasoning, and biologically relevant contexts, such as muscle work, organism movement, and energy transformation in body systems. Such an approach may help students develop a more meaningful and integrated understanding of work and energy as foundational concepts in both physics and biology.

REFERENCES

Adelia, I., Pranata, O. D., Seprianto, S., & Riana, P. C. (2024). Analisis Kesulitan Belajar Fisika Dasar Pada Mahasiswa Tadris Biologi IAIN Kerinci. EDU RESEARCH, 5(4), 647–654. https://doi.org/https://doi.org/10.47827/jer.v5i4.401

Aripiani, S. K., Samsudin, A., Kaniawati, I., Novia, H., Aminudin, A. H., Sutrisno, A. D., & Coştu, B. (2023). Diagnostic Instruments of Four-Tier Test Work and Energy (FORTUNE) to Identify The Level of Students' Conceptions. Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah, 8(1), 19–32. https://doi.org/10.24042/tadris.v8i1.13524

Banda, H. J., & Nzabahimana, J. (2021). Effect of integrating physics education technology simulations on students' conceptual understanding in physics: A review of literature. Physical Review Physics Education Research, 17(2), 23108. https://doi.org/10.1103/PhysRevPhysEducRes.17.023108

BOONE, W. J., TOWNSEND, J. S., & STAVER, J. R. (2016). Utilizing Multifaceted Rasch Measurement Through FACETS to Evaluate Science Education Data Sets Composed of Judges, Respondents, and Rating Scale Items: An Exemplar Utilizing the Elementary Science Teaching Analysis Matrix Instrument. Science Education, 100(2), 221–238. https://doi.org/https://doi.org/10.1002/sce.21210

Chabalengula, V. M., Sanders, M., & Mumba, F. (2012). DIAGNOSING STUDENTS' UNDERSTANDING OF ENERGY AND ITS RELATED CONCEPTS IN BIOLOGICAL CONTEXT. International Journal of Science and Mathematics Education, 10(2), 241–266. https://doi.org/10.1007/s10763-011-9291-2

Crouch, C. H., & Heller, K. (2014). Introductory physics in biological context: An approach to improve introductory physics for life science students. American Journal of Physics, 82(5), 378–386. https://doi.org/10.1119/1.4870079

Ding, L. (2023). Rasch Measurement in Discipline-Based Physics Education Research. In X. Liu & W. J. Boone (Eds.), Advances in Applications of Rasch Measurement in Science Education (pp. 19–46). Springer International Publishing. https://doi.org/10.1007/978-3-031-28776-3_2

Dyachenko, Ariana, Mukanova, Roza, & Erkibayeva, Meruert. (2024). Integration of interdisciplinary connections between chemistry, physics, and biology in the education of secondary school students. International Journal of Educational Reform, 10567879241290180. https://doi.org/10.1177/10567879241290180

Faihah, G., Sukarelawan, M. I., & Indratno, T. K. (2025a). Adaptation and Validation of a Four-Tier Misconception Diagnostic Test on Work and Energy Using Rasch Modeling. Jurnal Ilmiah Pendidikan Fisika, 9(3), 436–447. https://doi.org/https://doi.org/10.20527/jipf.v9i3.16781

Faihah, G., Sukarelawan, M. I., & Indratno, T. K. (2025b). Adaptation and Validation of a Four-Tier Misconception Diagnostic Test on Work and Energy Using Rasch Modeling. Jurnal Ilmiah Pendidikan Fisika, 9(3), 436–447. https://doi.org/https://doi.org/10.20527/jipf.v9i3.16781

Fatonah, U. (2024). Pengaruh efikasi diri dan motivasi siswa SMA terhadap pemahaman konsep pada materi usaha dan energi. Berkala Ilmiah Pendidikan Fisika. https://doi.org/https://doi.org/10.33394/jp.v10i2.7239

Fatonah, U., & Wicaksana, E. J. (2023). The Influence of Self-Efficacy and Learning Motivation of High School Students on The Understanding of The Concepts of Work and Energy. *EduFisika: Jurnal Pendidikan Fisika*, *8*(3), 325–335. https://doi.org/10.59052/edufisika.v8i3.29633

Firmansyah, D., Muhammad Reyza Arief Taqwa, Setiyani, A. ., & Intan Ramadani, C. (2023). Analysis of College Students' Conceptual Understanding on Work and Energy Topic in Various Representations. Jurnal Pendidikan Fisika Dan Teknologi, 9(2), 306–314. https://doi.org/10.29303/jpft.v9i2.4760

Hardini, I. P., Saregar, A., Subandi, Anugrah, A., Fitri, M. R., Panse, V. R., Umam, R., & Huda, S. (2026). Development of electronic student worksheet on disaster-based waves for eleventh-grade senior high school students. *AIP Conference Proceedings, 3226*(1), 020023. https://doi.org/10.1063/5.0329108

Hau, R. R. H., & Nuri, N. (2019). Pemahaman Siswa terhadap Konsep Hukum I Newton. Variabel, 2(2), 56–61. https://doi.org/http://dx.doi.org/10.26737/var.v2i2.1815

Hofer, S. I., Schumacher, R., & Rubin, H. (2017). The test of basic Mechanics Conceptual Understanding (bMCU): using Rasch analysis to develop and evaluate an efficient multiple choice test on Newton's mechanics. International Journal of STEM Education, 4(1), 18. https://doi.org/10.1186/s40594-017-0080-5

Hoskinson, A.-M., Couch, B. A., Zwickl, B. M., Hinko, K. A., & Caballero, M. D. (2014). Bridging physics and biology teaching through modeling. American Journal of Physics, 82(5), 434–441. https://doi.org/10.1119/1.4870502

Jewett Jr., J. W. (2008). Energy and the Confused Student I: Work. The Physics Teacher, 46(1), 38–43. https://doi.org/10.1119/1.2823999

Kaczmarek, M., & Greczyło, T. (2025). Challenges in Designing Basic Physics Course for Undergraduate Biologists - investigating difficulties with learning physics. Journal of Physics: Conference Series, 2950(1), 012013. https://doi.org/10.1088/1742-6596/2950/1/012013

Khusaini, K., Azizah, N., Suwasono, P., Yogihati, C. I., & Andriani, A. D. (2025). Rasch Model Application: Identification of Students' Conceptual Understanding of Static Fluid. Jurnal Pendidikan Fisika Indonesia, 21(1), 54–68. https://doi.org/https://doi.org/10.15294/jpfi.v21i1.15023

Laili, U. F., Huda, S., Anggraini, A., Maharani, F. C., Muniroh, L., & Umam, R. (2025). *Enhancing conceptual understanding and learning skills: The role of strategy, motivation, teacher support, and self-efficacy*. *Journal of Advanced Sciences and Mathematics Education, 5*(1), 27–41. https://doi.org/10.58524/jasme.v5i1.653

Laliyo, L. A. R., Sumintono, B., & Panigoro, C. (2022). Measuring changes in hydrolysis concept of students taught by inquiry model: stacking and racking analysis techniques in Rasch model. Heliyon, 8(3). doi: 10.1016/j.heliyon.2022.e09126

Lamanepa, G. H. (2021). Metode Pengukuran Rasch Dalam Pembelajaran Fisika. Jurnal Eduscience, 8(1), 47–52. https://doi.org/10.36987/jes.v8i1.2120

Lancor, R. A. (2014). Using Student-Generated Analogies to Investigate Conceptions of Energy: A multidisciplinary study. International Journal of Science Education, 36(1), 1–23. https://doi.org/10.1080/09500693.2012.714512

Lestari, F., Dalman, Noprisa, Efendi, D., Ariyanti, I., Anggara, B., & Umam, R. (2021). The effectiveness of math learning based on multiple intelligence: Implications at MTs Darul Ma'wa. *AIP Conference Proceedings, 2438*(1), 020001. https://doi.org/10.1063/5.0071296

Medvedev Oleg N. and Krägeloh, C. U. (2025). Rasch Measurement Model. In C. U. and S. R. J. and S. N. N. Medvedev Oleg N. and Krägeloh (Ed.), Handbook of Assessment in Mindfulness Research (pp. 131–147). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47219-0_4

Mulyani, S., Haniza, N., Ramadhani, D. G., & Mahardiani, L. (2021). Rash model approach for analysis of misconception on chemistry learning with distractor analysis. JKPK (Jurnal Kimia Dan Pendidikan Kimia), 6(1), 98–107. https://doi.org/10.20961/jkpk.v6i1.50517

Munifah, M., Tsani, I., Yasin, M., Tortop, H. S., Palupi, E. K., & Umam, R. (2019). Management system of education: Conceptual similarity (integration) between Japanese learning system and Islamic learning system in Indonesia. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah, 4*(2), 159–170. https://doi.org/10.24042/tadris.v4i2.4893

Nur'aini, D. A., Liliawati, W. ., & Novia, H. . (2023). The Effect of Differentiated Approach in Inquiry-based Learning on Senior High School Studentsâ€™ Conceptual Understanding of Work and Energy Topic. *Jurnal Penelitian Pendidikan IPA*, *9*(1), 117–125. https://doi.org/10.29303/jppipa.v9i1.2374

Opitz, S. T., Blankenstein, A., & Harms, U. (2017). Student conceptions about energy in biological contexts. Journal of Biological Education, 51(4), 427–440. https://doi.org/10.1080/00219266.2016.1257504

Opitz, S. T., Neumann, K., Bernholt, S., & Harms, U. (2019). Students' Energy Understanding Across Biology, Chemistry, and Physics Contexts. Research in Science Education, 49(2), 521–541. https://doi.org/10.1007/s11165-017-9632-4

Permatasari, A., Wartono, & Kusairi, S. (2018). Identification of students difficulties in terms of the higher order thinking skills on the subject of work and energy. AIP Conference Proceedings, 2014(1), 020052. https://doi.org/10.1063/1.5054456

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. Physical Review Physics Education Research, 15(2), 20111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Putra, F. G., Lengkana, D., Sutiarso, S., Nurhanurawati, N., Saregar, A., Diani, R., Widyawati, S., Suparman, S., Imama, K., & Umam, R. (2024). Mathematical representation: A bibliometric mapping of the research literature (2013–2022). *Infinity Journal, 13*(1), 1–26. https://doi.org/10.22460/infinity.v13i1.p1-26

Redish, E. F., Bauer, C., Carleton, K. L., Cooke, T. J., Cooper, M., Crouch, C. H., Dreyfus, B. W., Geller, B. D., Giannini, J., Gouvea, J. S., Klymkowsky, M. W., Losert, W., Moore, K., Presson, J., Sawtelle, V., Thompson, K. V, Turpen, C., & Zia, R. K. P. (2014). NEXUS/Physics: An interdisciplinary repurposing of physics for biologists. American Journal of Physics, 82(5), 368–377. https://doi.org/10.1119/1.4870386

Rivaldo, L., Taqwa, M. R. A., Zainuddin, A., & Faizah, R. (2020). Analysis of students' difficulties about work and energy. Journal of Physics: Conference Series, 1567(3), 032088. https://doi.org/10.1088/1742-6596/1567/3/032088

Samsudin, A., Cahyani, P. B., Rusdiana, D., Efendi, R., Aminudin, A. H., & Costu, B. (2021). Development of a Multitier Open-Ended Work and Energy Instrument (MOWEI) Using Rasch Analysis to Identify Students' Misconceptions. Cypriot Journal of Educational Sciences, 16(1), 16–32. https://doi.org/10.18844/cjes.v16i1.5504

Samsudin, A., Lestari, D., Kaniawati, I., Suhandi, A., Nasbey, H., Nurjannah, N., Aminudin, A. H., Sutrisno, A. D., Wibowo, F. C., & Prahani, B. K. (2025). Development of Work and Energy in Multi-tiers Test (WE-MusT) to identify students' conceptions levels: A Rasch analysis model. Kasetsart Journal of Social Sciences, 46(1), 460124.

Saregar, A., Putra, F. G., Diani, R., Anugrah, A., Misbah, M., & Umam, R. (2025). Innovative integrated disaster education in physics learning: An effort to enhance students' disaster literacy skills. *Jurnal Pendidikan IPA Indonesia, 14*(2), Article 23959. https://doi.org/10.15294/jpii.v14i2.23959

DOI: https://doi.org/10.30762/ijise.v5i2.8861

Schutte, L., Wissing, M. P., Ellis, S. M., Jose, P. E., & Vella-Brodrick, D. A. (2016). Rasch analysis of the meaning in life questionnaire among adults from South Africa, Australia, and New Zealand. Health and Quality of Life Outcomes, 14(1), 12. doi: 10.1186/s12955-016-0414-x.

Sekarningtias, F. O., Rusilowati, A., & Darsono, T. (2023). Tinjauan Literatur Sistematis: Tes Diagnostik untuk Mengidentifikasi Miskonsepsi Sains dengan Model Rasch. UPEJ Unnes Physics Education Journal, 12(3), 15–27. https://doi.org/https://doi.org/10.15294/upej.v13i1.72044

Sumintono, B., & Widhiarso, W. (2015). Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Trim Komunikata.

Taasoobshirazi, G., Bailey, M. L., & Farley, J. (2015). Physics Metacognition Inventory Part II: Confirmatory factor analysis and Rasch analysis. International Journal of Science Education, 37(17), 2769–2786. https://doi.org/10.1080/09500693.2015.1104425

Tong, D., Liu, J., Sun, Y., Liu, Q., Zhang, X., Pan, S., & Bao, L. (2023). Assessment of student knowledge integration in learning work and mechanical energy. Physical Review Physics Education Research, 19(1), 10127. https://doi.org/10.1103/PhysRevPhysEducRes.19.010127

Usman, M., Pramudawardani, H., Umam, R., & Takahashi, H. (2025). Synthesizing a framework and establishing content validity of an energy literacy instrument for Indonesian students. *Islamic Journal of Integrated Science Education (IJISE), 4*(3), 175–185. https://doi.org/10.30762/ijise.v4i3.6993