

Development of Two-Tier Test Instruments to Detect Student's Physics Misconception

Meilan Lengkong^{1*}, Edi Istiyono², B A O Rampean³, Awal Mulia Rejeki Tumanggor⁴, M F T Nirmala⁵

^{1,2} Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Sleman, Indonesia.

³ Master of Education in Chemistry, Graduate School, Universitas Negeri Yogyakarta, Sleman, Indonesia.

^{4,5} Physics Education, Graduate School, Universitas Negeri Yogyakarta, Sleman, Indonesia.

*Corresponding author. Email: lengkongmeilan@gmail.com

ABSTRACT

The aims of this research is to develop a two-tier test instrument to diagnose students' conceptual understanding abilities and knowing the level of student misconceptions on the topic of simple harmonic motion. The method in this research is 4D Research and Development (Define, Design, Develop, and Disseminate) by Thiagarajan and Semmel. The define stage is to conduct a literature study and observation in determining the focus of the problem. The design and develop stage is a modification of test instrument development procedure from the Oriondo & Dallo-Antonio models, namely: (1) Test design, (2) Trying out, (3) Test assembly, and measurement. The disseminate stage is conducting a seminar on the implementation result. The content validity of the nine items test was declared valid with Aiken's V coefficient ≥ 0.76 . The test instrument has been validated and declared valid and reliable has seen from the entire INFIT MNSQ item in the range of 0.77 - 1.30. The data analysis technique using the PCM 1 PL model with Polytomous scoring. The test instrument developed and tested to 60 students of Class X MIPA was able to measure the error of students by 30.5% having misconceptions and by 27.6% not understanding the concept. Of the nine items given with different conception, the largest percentage of students' misconceptions is in item number 4, which is 68.3% of students experiencing a misconception about the relationship of spring length to their frequency value.

Keywords: Development, diagnostic instruments, two-tier, misconception, simple harmonic motion.

1. INTRODUCTION

Learning is an effort to have knowledge and understanding, which is getting through a series of systematic learning processes, to form accurate, precise, and directed concepts. [1]. In learning science, students must actively think critically to strengthen their understanding of complex concepts in sciences, so that they can build knowledge and core ideas of the disciplines. [2], [3]. But there are still many factors that affect the lack of deepening of the concept of science, especially physics, one of the factors is Misconception [4], [5]. Because understanding a strong concept has an important role for students to develop their knowledge structure and to apply the right concepts in problem-

solving [5]–[7]. It is said to be a misconception because of the lack of correlation to actual opinions and know knowledge [8], [9]. If the aspects of knowledge and information provided to students differ from scientific theories coupled with applying personal gaze then it is clear that students experience misconceptions [10], [11]. knowledge [8], [9]. If the aspects of knowledge and information provided to students differ from scientific theories coupled with applying personal gaze then it is clear that students experience misconceptions [10], [11].

One of the materials with a high level of error in understanding is motion, especially about a simple harmonic motion which in its understanding requires analytical skills to link theory with phenomena or facts

[8], [12]. Some of the mistakes of students in understanding the concept, especially in the topic of simple harmonic motion such as the relation of the mathematical operational restoration force to the actual motion or phase angle [13], as well the students view that the magnitude of the amplitude depends on the period or the frequency. [14]. To overcome this, an instrument in the form of a test is needed to find out the students' understanding of the actual theory.

Diagnostic tests are deemed suitable for exposing scientific concept errors. The diagnostic test is an assessment instrument consisting of some questions to be tested, where the questions are focused on the difficulties and weaknesses of students in a material [15], [16]. One that is used in overcoming students' misconceptions is to use two-tier diagnostic formatted instruments [3]. Kahveci (2013) [17] states that two-tier items have the potential to elicit the way students choose reasons compared to using traditional multiple choice. Besides, the two-tier assessment is a practical method for daily use. Chandasegaran (2007) [18] states that the development of two-tier multiple choice diagnostic tests to identify alternative concepts of learners in a limited and clear area. Loh (2014) [19] states that two-tier multiple choice has a contribution to reducing the level of probability of answering correctly.

The two-tier diagnostic instrument is a multilevel assessment model in knowing the weakness and mistakes of students in understanding a concept. The first tier of this test instrument aims to measure students' ability to answer correctly questions consisting of five answer choices, and the second tier is the choice of reasons that refer to the first level answer [20], [21]. The use of the Two-tier test aims to reduce the guessing factor of students because students are required to provide the answers they choose [7]. Therefore, this study focuses on developing a systematic two-tier diagnostic test instrument to diagnose and detect students' understanding of concepts and misconceptions on simple harmonic motion material.

2. RESEARCH METHODS

This research refers to the 4D development stage (*Define, Design, Development, Disseminate*) according to Thiagarajan and Semmel [22] where is combined with the development stage of the modified Oriondo and Antonio model test instrument were consists of three-stage, namely: (1) Design the test, (2) Trial test, (3) assembly and measurement tests [23]. This research aims to diagnose students' abilities and to determine the level of student misconceptions as measured through the

development of a two-tier test instrument. Literature and observation studies were doing at the *define* stage. In the *design* procedure, design tests and trials carried out. The design of the test includes developing test specifications, Writing tests, Reviewing tests [1]. The testing stage includes testing to determine validity and reliability, then carried out improvements. The *development* stage is Assemble tests and measuring include assembling tests, carrying out a test, and interpreting test results, in this case, is to measure students' misconceptions in simple harmonic motion material. The next stage is *disseminating*, the test instruments that have been assembling and carried out measurements at the seminar to be follow up.

2.1. Population and sampling

This research was conduct in senior high school in the Yogyakarta from February to March 2020. This research using a simple random sampling technique. The empirical test has been doing in 113 students of class XI who had studied the simple harmonic motion material, consisting of two classes in SMA Negeri 1 Banguntapan and two classes in SMA Negeri 1 Sewon. After obtaining a valid and reliable instrument, then a field test was carried out, namely the measurement in class X involving 60 students as subjects.

2.2. Data Analysis

The data analysis techniques are using Quest and Parscale program. The data obtained were polytomous data using four categories and were analyzed based on Partial Credit Model 1-PL (PCM 1-PL).

Two-tier test instrument, at the first tier the test item has 5 choices of answers A, B, C, D, E based on the concept. The second tier contains 5 answers in the form of reasons for choosing answers at the first tier [16]. The results of students' answers to each item will provide a combination of the answers presented in the Table 1.

Table 1. decision criteria from a combination of answers

No.	Answer (1 st – tier)	Scientific reason (2 nd – tier)	Decision category	Score
1	True	True	Understand the concept (UC)	4
2	False	True	Misconception (MSC)	3
3	True	False	Misconception (MSC)	2
4	False	False	Not understand the concept (NC)	1

3. RESULT AND DISCUSSION

The Define stage carried out literature studies and observations related to problems that occurred in the field, especially the difficulties and weaknesses of students in learning physics. The information obtained through research at this stage then determines the focus of the problem to be investigated and to find a solution, one of which is the simple harmonic motion physics material where students experience many misconceptions.

Furthermore, the stages of research must be carried out in accordance with the explanation of the research method *design* and *develop*. From the results of the design, it can be seen that the instrument item matches, instrument validity, item difficulty level, reliability, total information function, and SEM. After the instrument is declared suitable, then the test instrument is tested to find out the diagnosis results of students' misconceptions. Content validity obtained from assessments conducted by 7 raters with 4 rating scales. The minimum value of Aiken's V Coefficient based on the V-Aiken table is 0.76 and a probability of 0.045. The results of the content validity for the 9 test items showed a coefficient of V-Aiken ≥ 0.76 . Based on the analysis of content validity, the data obtained from the item items categories shown in Table 2.

Table 2. Analysis Content Validity based on Aiken's V

Item Number	V-Aiken Coefficient	Category
1	0,77	Valid
2	0,76	Valid
3	0,76	Valid
4	0,76	Valid
5	0,77	Valid
6	0,76	Valid
7	0,76	Valid
8	0,76	Valid
9	0,76	Valid

3.1 Item Reliability

The item reliability value can know from the summary of the item estimate shows the number is 0.73, and the reliability of people is based on the case estimate is 0.70.

3.2 Item Estimation

At the development stage, nine test items that proved valid were then tested and measured on 113 students. Student responses to the developed test items were

analyzed based on the Partial Credit Model (PCM) PCM 1 PL. The results of the test compatibility analysis based on empirical data can see in Table 3.

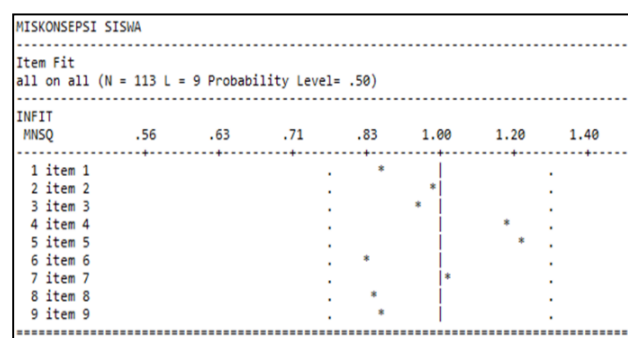
Table 3. Test fit Statistical parameters at 0,50 probability level

Aspect	Item Estimation	Case Estimation
Average value and standard deviation	$0.01 \pm 0,81$	0.26 ± 0.73
Adjusted average and standard deviation	0.69	0.61
The average value and INFIT MNSQ standard deviation	0.97 ± 0.14	0.97 ± 0.56
The average value and the OUTFIT MNSQ standard deviation	1.14 ± 0.56	1.14 ± 1.60
Average difficulty	0.00 ± 0.81	

3.3 Item Validity

Polytomous data with 4 scoring criteria as analyze with the Rasch Model. Test items analyzed by PCM 1-PL analysis results with the Quest program on 9 test items give INFIT MNSQ Output results depending on the range of 0.77 to 1.30 means all items fit of the PCM model [2], [24]. The results of the validity test show the compatibility of the test items with the Rasch Model shown in Figure 1.

Figure 1. Item Distribution with the Rasch model



3.4 Item Difficulty

Item difficulty Index can be seen through analysis with the Quest program presented in Diagram in Figure 2.

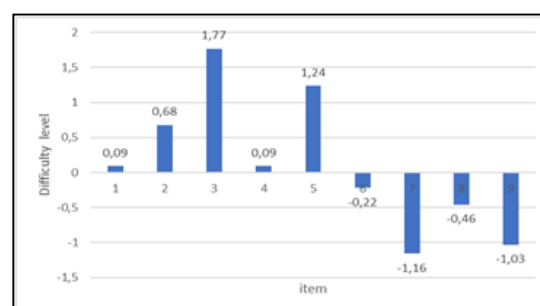


Figure 2. Graph of items difficulty level

Based on Figure 2 the level of difficulty items has difficult, medium, and easy categories. The items are categorized as difficult if they have a difficulty index of 0.00 - 0.30, are categorized as moderate with a difficulty index range of 0.31 - 0.70, and said to be easy with a range of 0.71 - 1.00. The greater the difficulty item index obtained from the calculation results, the easier the problem is. [1] The degree of appropriateness of items can affect the total form of the test score. For tests that are very difficult to make at a score < 0.25 positive skewed distribution form. While items with easy difficulty with a score of > 0.80 have a negatively skewed distribution [1]. All the items in this diagnostic instrument are categorized as good because they are in the range of scores from -2.0 to +2.0 [25].

3.5 Information function and Standard error measurement

From the results of item analysis with the Parscale program, the Total Information Function (TIF) is obtaining. The reliability of the test instrument based on the item response theory can see through the Total Information Curve (TIC) graph and the Standard Error of Measurement (SEM) in Figure 3.

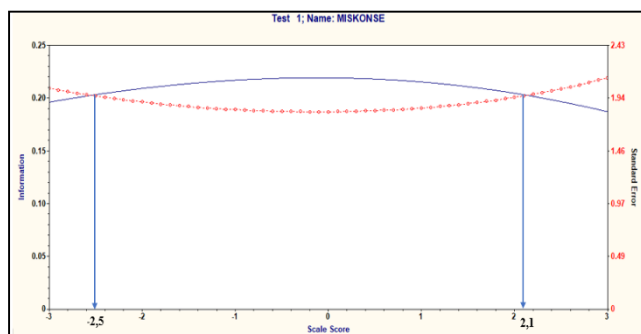


Figure 3. Information function and SEM

Figure 3 shows the TIF and SEM curves in evaluating and measuring the ability to understand the concepts of simple harmonic physics material. The information function value gives a value of 0.2 wherein the graph is shown through the intersection line on the left scale. While for SEM value, the result is 1.94, which is seen in the intersection line which is the right scale on the graph. From this curve, it is known that the instrument for measuring students' misconceptions is more precisely tested on respondents who have a minimum ability of -2.5 and a high ability of +2.1.

This wide range of abilities is very appropriate to be applied in high schools, considering that the input of students has a variety of backgrounds and abilities. This can be seen through the Histogram in Figure 4.

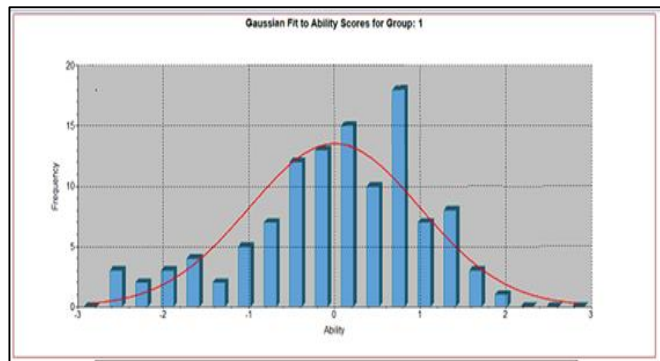


Figure 4. Distribution of student misconception ability

Students' abilities can be identified through the output analysis using the Parscale program. Students' conceptual understanding abilities are presented on a logit scale in the ability column [2], [24]. The results of the measurement on the 60 students gave scores that were in the ability range -3 to +3.

3.6 Diagnosis of student misconceptions

The validated test instrument was then developed by completing the two-tier test format in diagnosing specific student misconceptions on each simple harmonic motion sub-topic. Data analysis is presented in Table 4.

Table 4. Percentage of misconceptions in each sub-topic

The measured sub-topic	Number of student response categories			Percentage of misconceptions (%)
	UC	MSC	NC	
The pendulum period is perpendicular to the length of the rope	18	29	13	48,3
The pendulum period is influenced by the mass of the object that is swinging	20	2	38	3,3
Frequency in simple harmonic motion	6	32	22	53,3
The frequency is perpendicular to the change in spring length	12	41	7	68,3
The spring constant is inversely proportional to the frequency	3	19	19	31,6
The direction of the swing restoration force is in the direction of the amount of force applied	35	14	11	23,3
The period of vibration is caused by the recovery force	48	8	4	13,3
The direction of the restoring force on the spring always approaches the deviant force	39	10	11	16,6
The amount of deviation is not affected by the movement in the spring.	45	10	5	16,6
Percentage	41,9	30,5	27,6	564

The instrument for diagnosing student misconceptions was distributed to 60 students of class X MIPA program consisting of 9 items. Based on the analysis using the decision making a principle of two-tier tests, the results obtained by 41.9% of students understand the concept of simple harmonic motion. This happens due to several factors, one of them is the difficulty of the items as evidenced by the achievement of students' understanding of the concept in point seven, which is 80% understanding the concept, although the difficulty level of the seventh item is categorized as difficult with negative skewed distribution, which means the tail of the distribution is on the left, which shows that most of the values are on the right-hand side of the curve. The analysis also showed 30.5% of students experienced misconceptions. Students agree to blame because they make the first two questions and the second chooses one and the second choice of answers is one of the correct reasons. In this analysis, it can be determined the percentage criteria for students' misconceptions for each sub-topic being tested. The highest percentage of misconceptions on item 68.4% of students answered misconceptions about the relationship of spring length to the frequency value. This result is in line with research conducted on 45 students who carried out special classes at the Chiang Mai University of Thailand and provided information that most students reasoned that the mass and spring constant are the same so that the frequency value must remain the same [26]. Students' misconceptions in the medium category are shown in items 1, 3, and 5, and low misconception in items 2, 6, 7, 8, and 9 [27]. Meanwhile, 27.6% as a whole did not understand the concept and when viewed from the sub-topic of 63.3% students did not understand or did not know the correct concept in the second item which when viewed from the difficulty level of the item was categorized as being to measure the ability students to the concept of the mass of objects that swing affects the pendulum period.

The form of reason presented in this two-tier test instrument is a closed reason to minimize students who only guess the answer at the first tier. So that it can be known in detail the extent to which students can understand the concepts in a material.

4. CONCLUSION

From the results of research and development of instruments to diagnose students' misconceptions on the topic of simple harmonic motion, they have been testing for validity and reliability using a two-tier test format that is in the proper and good category according to the V-Aiken coefficient ≥ 0.76 . The instrument has been

validated and declared valid and reliable has seen from the entire INFIT MNSQ item in the range of 0.77 - 1.30. The test instrument develops to measure students' misconceptions that useful for knowing students' abilities to understand of physics concepts based on the categorization of students' answers in understanding concepts correct, so that students who experience misconceptions and students who do not understand concepts can be known.

AUTHORS' CONTRIBUTIONS

Meilan lengkong analyzed the preliminary studies, developed two-tier instrument, participated in the research design, and performed the statistical analyses. E Istiyono participated in the sequence alignment and coordinated the manuscript. B A O Rampean participated in the design of study. A M R Tumanggor performed the statistical analyses, sequence alignment and helped to draft the manuscript. M F T Nirmala drafted the manuscript and helped in the design of the study. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The author would like to thank all respondents who participated in this research and to the Physics Teachers of SMA N 1 Sewon and SMA N 2 Banguntapan are have supported the implementation of this research.

REFERENCES

- [1] E. Istiyono, *Pengembangan Instrumen Penilaian dan Analisis Hasil Belajar Fisika*, 1st ed. Yogyakarta: UNY Press, 2018.
- [2] E. Istiyono, W. Brams Dwandaru, and F. Dan Rahayu, "The developing of creative thinking skills test based on modern test theory in physics of senior high schools," *Cakrawala Pendidik.*, vol. 37, no. 2, pp. 190–200, 2018, doi: 10.21831/cp.v37i2.19233.
- [3] Y. Xiao, J. Han, K. Koenig, J. Xiong, and L. Bao, "Multilevel Rasch modeling of two-Tier multiple choice test: A case study using Lawson's classroom test of scientific reasoning," *Phys. Rev. Phys. Educ. Res.*, 2018, doi: 10.1103/PhysRevPhysEducRes.14.020104.
- [4] Z. D. Kirbulut and O. Geban, "Using three-tier diagnostic test to assess students' misconceptions of states of matter," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 10, no. 5, pp. 509–521, 2014, doi: 10.12973/eurasia.2014.1128a.
- [5] G. Liu and N. Fang, "Student misconceptions about force and acceleration in physics and engineering mechanics education," *Int. J. Eng. Educ.*, vol. 32, no. 1, pp. 19–29, 2016.

- [6] A. M. R. Tumanggor, J. Jumadi, I. Wilujeng, and E. S. Ringo, "The Profile of Students' Physics Problem Solving Ability in Optical Instruments," *J. Penelit. Pengemb. Pendidik. Fis.*, vol. 5, no. 1, pp. 29–40, 2019, doi: 10.21009/1.05104.
- [7] R. A. Streveler, T. A. Litzinger, R. L. Miller, and P. S. Steif, "In the engineering sciences: Overview and future research directions," *J. Eng. Educ.*, no. July, pp. 279–294, 2008.
- [8] I. Kuczmann, "The structure of knowledge and students' misconceptions in physics," *AIP Conf. Proc.*, vol. 1916, no. December, 2017, doi: 10.1063/1.5017454.
- [9] A. Grigorovitch, "Children's misconceptions and conceptual change in Physics Education: the concept of light," *J. Adv. Nat. Sci.*, vol. 1, no. 1, pp. 34–39, 2014, doi: 10.24297/jns.v1i1.5037.
- [10] A. M. R. Tumanggor, Supahar, H. Kuswanto, and E. S. Ringo, "Using four-tier diagnostic test instruments to detect physics teacher candidates' misconceptions: Case of mechanical wave concepts," *J. Phys. Conf. Ser.*, vol. 1440, no. 1, 2020, doi: 10.1088/1742-6596/1440/1/012059.
- [11] M. Bozzi *et al.*, "Highlight Misconceptions in Physics: a T.I.M.E. Project," *INTED2019 Proc.*, vol. 1, no. March, pp. 2520–2525, 2019, doi: 10.21125/inted.2019.0689.
- [12] A. M. R. Tumanggor, Supahar, E. S. Ringo, and M. D. Harliadi, "Detecting Students' Misconception in Simple Harmonic Motion Concepts Using Four-Tier Diagnostic Test Instruments," *J. Ilm. Pendidik. Fis. Al-Biruni*, vol. 9, no. 1, pp. 21–31, 2020, doi: 10.24042/jipfalbiruni.v9i1.4571.
- [13] S. Somroob and P. Wattanakasiwich, "Investigating student understanding of simple harmonic motion," *J. Phys. Conf. Ser.*, vol. 901, no. 1, 2017, doi: 10.1088/1742-6596/901/1/012123.
- [14] D. A. Nugraha, C. Cari, A. Suparmi, and W. Sunarno, "Physics students' answer on simple harmonic motion," *J. Phys. Conf. Ser.*, vol. 1153, no. 1, 2019, doi: 10.1088/1742-6596/1153/1/012151.
- [15] L. Halim, T. K. Yong, and T. S. M. Meerah, "Overcoming Students' Misconceptions on Forces in Equilibrium: An Action Research Study," *Creat. Educ.*, vol. 05, no. 11, pp. 1032–1042, 2014, doi: 10.4236/ce.2014.511117.
- [16] D. K. Gurel, A. Eryilmaz, and L. C. McDermott, "A review and comparison of diagnostic instruments to identify students' misconceptions in science," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 11, no. 5, pp. 989–1008, 2015, doi: 10.12973/eurasia.2015.1369a.
- [17] Kahveci, *Concepts of Matter in Science Education: Part II*. Dordrecht: Springer, 2013.
- [18] A. L. Chandrasegaran, D. F. Treagust, and M. Mocerino, "The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation," *Chem. Educ. Res. Pract.*, vol. 8, no. 3, pp. 293–307, 2007, doi: 10.1039/B7RP90006F.
- [19] A. S. L. Loh, R. Subramaniam, and K. C. D. Tan, "Exploring students' understanding of electrochemical cells using an enhanced two-tier diagnostic instrument," *Res. Sci. Technol. Educ.*, vol. 32, no. 3, pp. 229–250, 2014, doi: 10.1080/02635143.2014.916669.
- [20] B. K. Bayrak, "Using Two-Tier Test to Identify Primary Students' Conceptual Understanding and Alternative Conceptions in Acid Base," *Mevlana Int. J. Educ.*, vol. 3, no. 2, pp. 19–26, 2013, doi: 10.13054/mije.13.21.3.2.
- [21] C. Y. Tsui and D. Treagust, "Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument," *Int. J. Sci. Educ.*, vol. 32, no. 8, pp. 1073–1098, 2010, doi: 10.1080/09500690902951429.
- [22] Thiagarajan and Sivasailan, "Instructional development for training teachers of exceptional children: A sourcebook," *J. Sch. Psychol.*, vol. 14, no. 1, p. 75, 1976, doi: 10.1016/0022-4405(76)90066-2.
- [23] Oriondo and Dallo-Antonio, *Evaluating Educational outcomes (test, measurement, and evaluation)*, 5th editio. Queson city: REX Printing Company, Inc, 1984.
- [24] S. Supahar and Z. K. Prasetyo, "Pengembangan Instrumen Penilaian Kinerja Kemampuan Inkuiri Peserta Didik Pada Mata Pelajaran Fisika SMA," *J. Penelit. dan Eval. Pendidik.*, vol. 19, no. 1, pp. 96–108, 2015, doi: 10.21831/pep.v19i1.4560.
- [25] R. K. Hambleton, H. Swaminathan, and H. J. Rogers, *Fundamentals of Item Response Theory Library*. SAGE Publications. Inc, 1991.
- [26] N. Tongnopparat, J. Poonyawatpornkul, and P. Wattanakasiwich, "Investigation of Student Reasoning about Harmonic Motions," vol. 017033, pp. 80–83, 2014, doi: 10.7566/jpscp.1.017033.
- [27] Y. Kurniawan, "Investigation of the Misconception in Newton II Law," *J. Pena Sains*, vol. 5, no. 1, p. 11, 2018, doi: 10.21107/jps.v5i1.3879.