

## DEVELOPMENT FOUR-TIER DIAGNOSTIC TEST INSTRUMENTS ON NEWTON'S LAW OF GRAVITATION CONCEPT

Nurul Izzah and Madlazim

Physics Department, Faculty of Mathematics and Natural Science, State University of Surabaya  
Email: nurulizzah@mhs.unesa.ac.id

### Abstract

Misconception is knowledge or idea that not in accordance with the scientific knowledge or not agreed upon by scientists. Misconceptions can be found on the concepts of physics, including Newton's Law of Gravitation concept. This paper aim is to develop a four-tier diagnostic test instrument for Newton's Law of Gravitation concept. This developed instrument is amount to 18 questions which contained concept, including sub material Mass and Weight, Gravitational Field, etc. The type of this research is development research using the ADDIE model. The result of the calculation shows that the developed instrument was declared valid by two validator lecturers with a validity level between (83–95) %. The result of the instrument's external validity which was tested on 20 students was obtained 15 questions valid and reliable.

**Keywords:** four-tier diagnostic test, misconceptions, Newton's Law of Gravitation

### Abstrak

Miskonsepsi merupakan pengetahuan atau gagasan yang salah dan tidak sesuai dengan pengetahuan ilmiah ataupun tidak disepakati oleh para ilmuwan. Miskonsepsi dapat ditemukan pada konsep-konsep Fisika, termasuk materi Hukum Gravitasi Newton. Penelitian ini bertujuan mengembangkan instrumen *four-tier diagnostic test* untuk materi Hukum Gravitasi Newton. Instrumen yang dikembangkan berjumlah 18 pertanyaan yang berisi soal-soal konsep meliputi sub materi Massa dan Berat, Gaya Gravitasi, dll. Jenis penelitian ini merupakan penelitian pengembangan dengan menggunakan model ADDIE. Berdasarkan hasil perhitungan, instrumen dinyatakan valid oleh dua dosen validator dengan tingkat validitas antara (83–95) %. Hasil uji coba instrumen validitas eksternal yang dilakukan kepada 20 peserta didik diperoleh 15 butir soal valid dan reliabel.

**Kata kunci:** *four-tier diagnostic test*, miskonsepsi, Hukum Gravitasi Newton

### INTRODUCTION

Physics is a branch of science that studies natural phenomena, objects in nature, and interactions between these objects. Some of the physics materials can be learned from experiences in everyday life. Other physics materials discuss something that cannot be displayed in the classroom and abstract, like Newton's Law of Gravitation. That's because the main discussion of Newton's Law of Gravitation material includes objects that are very large ( $\geq 10^{23}$  kg) (Giancoli, 2014) and only a small portion can be seen clearly, like the Earth, the Sun and other planets. The abstract discussion of Newton's Law of Gravitation can potentially cause misconception of the learners (Fauzia & Madlazim, 2015),

Concept is the definition or characteristic of something that is universally agreed, while the students'

interpretation of a concept presented by Teacher called conception (Wiyono, et al., 2016). The difference between the conception of the concepts that have been agreed universally cause someone having misconceptions. The misconception can happen to any concept. One example of misconceptions that happened on Newton's Law of Gravitation concept is students consider that the shape of the planet's orbit around the sun is oval and the Sun is at the center of the oval. Other students consider the shape of the planet's orbit is a circle with the Sun is at the center. The correct concept is the planet's orbit around the Sun shaped oval with the Sun is at one focus point.

The misconception that if maintained would potentially hinder students to understand the next material related to the concepts that have been studied,

therefore the students' misconceptions need to be detected early. There are several methods that can be used to detect misconceptions. According to Kirbulut & Geban (2014), one of the methods that can be used to determine the students' misconceptions effectively is diagnostic tests. A diagnostic test is a test used to determine the thinking of students in answering the questions and the reasons stated.

One type of effective diagnostic tests to identify students misconceptions is a four-tier diagnostic test. The four-tier diagnostic test instrument is a multiple-choice test that consists of four levels. At the first level is choice of answers of questions, the second level is a choice of answers confidence, the third level contains the reasons of the answers, and the fourth level is the choice of reasons confidence. Standard good instruments by Sugiyono (2009) is an instrument that fulfills the aspect of validity and reliability. Validity indicates the accuracy of an instrument to use appropriate measuring function (Matondang, 2009). Reliability is a coefficient which shows the consistency of the instrument when used to measure the same thing repeatedly (Khumaedi, 2012: 26),

Based on the above description, the Researcher necessary to develop misconceptions diagnostic test instruments with the four-tier type of Newton's Law of Gravitation concept. The purpose of this study was to determine the feasibility of the instrument in terms of aspects of validity and reliability aspects.

**METHOD**

This research type is the development research by using ADDIE models (Analyze, Design, Development, Implementation, and Evaluation), adapted from Anglada (2007). The resulting product from this research is the diagnostic test instruments with four-tier type to identify misconceptions students on Newton's Law of Gravitation concept which amount to 18 questions.

The process to produce the instrument begins with the initial trial with the 30 students of Physics Unesa to determine answer choice and reasons choice. Answers and reasons gathered from initial trials and then be used as a multiple-choice on the developed instruments. The developed instrument further explored and validation internally by two expert lecturers to determine the validity of the instrument in the aspect of content, construct, and language. Internal validity is obtained by calculating the percentage of the validity in every aspect. Internal validity of instruments fulfilled if the average percentage of validity is less than 61% (Riduwan & Akdon, 2013).

Having obtained the internal validity, the instrument was tested in limited to 20 students in class XI SMAN 1

Gresik to determine the content empirical validity, construct empirical validity and reliability of the instrument. The instrument stated to fulfill the content validity if the percentage of false positive (FP) and false negative (FN) is less than 10% (Hestenes & Halloun, 1995). Construct empirical validity obtained by calculating the product moment correlation coefficient ( $r_{xy}$ ) of the instrument and the question items. According to Arikunto (2016). If the value of the product moment correlation coefficient ( $r_{xy}$ ) greater than the reference value ( $r_{table}$ ), the developed instrument and the question items can be declared valid. The reference value ( $r_{table}$ ) in this study was 0.444 (Sugiyono, 2015).

From the limited test results data, can also be obtained the reliability of the instrument. The reliability can be determined by calculating Alpha Cronbach's coefficient. The instrument can be declared reliable if the value of  $r_{xy} > r_{tabel}$  (Arikunto, 2016).

**RESULTS AND DISCUSSION**

The initial activity of this research is doing the initial trial to 30 students of Physics Unesa to determine the answers choice and choice of potential reasons in accordance with the thinking of students. The next step is to construct the four-tier diagnostic test instrument using the results of multiple choice answers on the initial trial. Then the developed instrument analyzed and validated internally by two lecturers of Physics to determine the validity of the instrument. The recapitulation of internal validity by two expert lecturers can be seen in Table 1.

**Table 1.** Recapitulation Internal Validity Assessment

Aspect	Percentage (%)	Validity Criteria
Contents	84	Very Valid
Construct	95	Very Valid
Language	93	Very Valid

Based on Table 1, it is known that the developed instrument can be declared valid internally with very valid criteria.

Instrument *four-tier diagnostic test* which has been declared valid internally subsequently tested limited to 20 students XI grade in SMAN 1 Gresik to determine the feasibility of external instrument which includes the content empirical validity, construct empirical validity, and reliability of the instrument. The analysis results of the contents empirical validity recapitulation presented in Table 2 below.

**Table 2.** Content Empirical Validation Results

No. Question	false Positive (FP)	false Negative (FN)
1	7	1
2	4	0
3	1	3
4	3	1
5	0	1
6	0	0
7	0	0
8	2	0
9	0	0
10	0	0
11	0	0
12	5	0
13	1	0
14	4	0
15	3	2
16	0	1
17	1	0
18	0	0
total	31	9
Percentage	8.61%	to 2.50%

Based on Table 2, it can be seen that the value of FP and FN is 8.61% and 2.50%, therefore the four-tier diagnostic test instrument on Newton's Law of Gravitation concept that has been developed fulfill the content empirical validity. According to Hestenes & Halloun (1995) the instrument will fulfill the content validity if the FP and FN values are less than 10%. This result is also consistent with the research by Kaltakçi, Eryilmaz, & McDermott (2017) entitled "Development and Application of A Four-Tier Test to Assess Pre-Service Physics Teachers' Misconceptions about Geometrical Optics" in ERIC, the results of these studies shows the percentage of FP and FN are 3.5% and 3.3%.

Then the limited trial results used to determine the construct empirical validity of instruments. The value of the product moment correlation coefficient of the instrument is obtained 0.610 and the reference value is 0.444. Thus, the developed instrument can be said to fulfill the empirical validity of the construct because  $r_{xy} > r_{table}$  (Arikunto, 2016: 89). The criteria of construct empirical validity of instruments is high.

After knowing the value of the construct empirical validity instrument, it was continued by testing the empirical validity of question items. The results of construct empirical validity of each question item can be seen in Table 3 below.

**Table 3.** Empirical Results Construct Validity Grain Problem

No. Question	$r_{xy}$	$r_{table}$	Criteria	Category
1	0.711	0.444	valid	High
2	0.860		valid	Very high
3	0.477		valid	Medium
4	0.456		valid	Medium

No. Question	$r_{xy}$	$r_{table}$	Criteria	Category
5	-0.160	0.444	Invalid	Invalid
6	0.456		valid	Medium
7	0.711		valid	High
8	0.516		valid	Medium
9	0.629		valid	High
10	0.472		valid	Medium
11	0.711		valid	High
12	0.711		valid	High
13	0.537		valid	Medium
14	0.283		Invalid	Low
15	0.445		valid	Medium
16	0.711		valid	High
17	0.224		Invalid	Low
18	0.623		valid	High

Based on Table 3, shows that of the 18 questions that have been developed, there are 3 questions that do not fulfill the construct empirical validity because  $r_{xy}$  values that are less than the  $r_{table}$  value. Thus, only 15 items were declared valid and may be used in the reliability test.

From the calculation of the reliability test instrument, is obtained Alpha Cronbach coefficient values for 0.843. The coefficient value is greater than the  $r_{table}$  value and can be stated that the developed instrument is feasible and reliable. (Arikunto, 2016)

## CLOSING

### Conclusion

Based on the analysis and discussion of research data obtained, it can be concluded that the four-tier diagnostic test instrument on Newton's Law of Gravitation concept that has been developed is feasible because it fulfills the aspect of validity and reliability of the instrument.

### Suggestion

Based on the research that has been done, there are some suggestions for further research as follows:

1. Doing further research to reduce misconceptions on Newton's Law of Gravitation concept using a four-tier diagnostic test instrument that has been developed.
2. To develop a four-tier instrument diagnostic test for other physics concept to identify misconceptions that occur on the students.

## REFERENCES

- Anglada, D. (2007). *An Introduction to Instructional Design: Utilizing a Basic Design Model*.
- Arikunto, S. (2016). *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: PT. Bumi Aksara.

- Fauzia, L. A., & Madlazim. (2015). Pengaruh Penerapan Model Pembelajaran Diskusi Kelas terhadap Penurunan Miskonsepsi Siswa pada Pembelajaran Fisika SMA Materi IPBA. *Jurnal Inovasi Pendidikan Fisika (JIPF)*, 04(01), 26–29.
- Giancoli, D. C. (2014). *Physics Principles with Application* (Seventh). United States of America: Pearson Education, Inc.
- Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory. *The Physics Teacher*, 502–506.
- Kaltakçi, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and Application of a Four-Tier Test to Assess Pre-Service Physics Teachers' Misconception about Geometrical Optics. ERIC, 35(Research in Science and Technological Education), 238–260.
- Khumaedi, M. (2012). Reliabilitas Instrumen Penelitian Pendidikan. *Jurnal Pendidikan Teknik Mesin*, 12(1), 25–30.
- Kirbulut, Z. D., & Geban, O. (2014). Using Three-Tier Diagnostic Test to Assess Students' Misconceptions of States of Matter. *Eurasia Journal of Mathematics, Science and Technology Education*, 10(5), 509–521. <https://doi.org/10.12973/eurasia.2014.1128a>
- Matondang, Z. (2009). Validitas dan Reliabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa PPS Unimed*, 6(1), 87–97. <https://doi.org/10.1017/CBO9781107415324.004>
- Riduwan, & Akdon. (2013). *Rumus dan Data dalam Analisis Statistika*. Bandung: Alfabeta.
- Sugiyono. (2009). *Metode Penelitian Administrasi Dilengkapi dengan Metode R&D*. Bandung: Alfabeta.
- Sugiyono. (2015). *Metode Penelitian Pendidikan*. Bandung: Alfabeta.
- Wiyono, F. M., Sugiyanto, & Yulianti, E. (2016). Identifikasi Hasil Analisis Miskonsepsi Gerak Menggunakan Instrumen Diagnostik Three Tier pada Siswa SMP. *Jurnal Penelitian Fisika Dan Aplikasinya (JPFA)*, 06(02), 61–69.