

## THE VALIDITY OF FOUR-TIER'S MISCONCEPTION DIAGNOSTIC TEST FOR DYNAMIC FLUID CONCEPTS

**Dyah M. Kurniawati dan Frida U. Ermawati**

Physics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Surabaya

Email: dyahkurniawati1@mhs.unesa.ac.id

frida.ermawati@unesa.ac.id

### Abstract

Physics was considered by students as the most difficult lesson. One of the reasons is that Physics contains many formulas without being followed by a good explanation of the Physics concepts behind the formulas from the teacher. Such situation potentially causes misconceptions to students. This paper reports the work to develop a four-tier misconception diagnostic test instrument in Dynamic Fluid concepts. The aim was to determine the feasibility of the developed instrument in terms of validity and reliability. The instrument's validation result was obtained after two independent Physics Lecturers evaluated the content of the instrument. The reliability results was gathered from the test trial to 50 numbers of 11 grade students in Science class of senior high school 1 Sidoarjo. The results show that the validity was 92.55 %. The false positive and false negative values were 4.3% and 3.8%, respectively which means that the instrument is valid because the values less than 10 %. The correlation coefficient of Pearson product moment ( $r_{xy}$ ) was 0.591 while the reliability index value ( $r_{11}$ ) was 0.751. Those values are greater than the  $r_{table}$  of 0.279, which means that the developed instrument is reliable.

**Keywords:** Dynamic fluid concepts, four-tier diagnostic test, validity, reliability

### Abstrak

Fisika dianggap oleh siswa sebagai pelajaran yang paling sulit. Salah satu alasannya adalah bahwa Fisika mengandung banyak rumus tanpa diikuti oleh penjelasan mengenai konsep fisika dibalik rumus tersebut dari guru dengan baik. Situasi seperti itu berpotensi menyebabkan kesalahpahaman dalam siswa. Makalah ini melaporkan pekerjaan untuk mengembangkan instrumen uji diagnostik miskonsepsi empat tingkat dalam konsep Fluida Dinamik. Tujuannya adalah untuk menentukan kelayakan instrumen yang dikembangkan dalam hal validitas dan reliabilitas. Hasil validasi instrumen diperoleh setelah dua Dosen Fisika independen mengevaluasi konten instrumen. Reliabilitas dikumpulkan dari uji coba ke 50 jumlah siswa kelas 11 di kelas Sains SMA 1 Sidoarjo. Hasil penelitian menunjukkan bahwa validitasnya adalah 92,55%. Nilai *false positives* dan *false negatives* masing-masing adalah 4,3% dan 3,8%, yang berarti bahwa instrumen tersebut valid karena kedua nilai tersebut jauh di bawah 10 %. Koefisien korelasi *Pearson product moment* ( $r_{xy}$ ) adalah 0,591 sedangkan nilai indeks reliabilitas ( $r_{11}$ ) adalah 0,751. Nilai-nilai tersebut lebih besar dari  $r_{table}$  0,279, yang berarti bahwa instrumen yang dikembangkan dapat diandalkan

**Kata kunci:** Fluida dinamik, tes diagnostik berformat *four tier*, validitas, reliabilitas

### INTRODUCTION

Physics is considered by students as the difficult subject. The reason is that Physics subject contains many formulas, explanations on the physics concepts behind the formulas given by the teachers are also very limited and unclear hence less understandable. These situations caused students to be reluctant to repeat the lessons at home. Besides, the evaluation test scores in Physics subject earned by students are never satisfying. If such bad

learning atmosphere took place continuously, this potentially may cause disruption of students' understanding on the following concepts.

In daily life, students has gained initial knowledge on Physics that comes from their interaction, their environment to their family members as well as to the friends in the neighbourhood. For example, students have the initial knowledge in accordance with the Bernoulli's Law, namely on mosquito sprayers. When someone wants

to kill the mosquitos, what they do is to pump the mosquito spray by pulling and pressing the piston rod alternately so that the mosquito repellent liquid inside the mosquito spray container will move up through the capillary tube inside the container and burst out to fill the room. The effect if the mosquitos are found in that room, they inhaled the poisonous spray and the mosquitos die. If for a certain reason the capillary tube hole is blocked by something (for example by dust), then someone can insert a thin wooden stick (a toothpick stick) into the blocked tube and move it up and down quickly. By this way, the blocking is reopened. Students may not be aware that the mosquito sprayer works based on the Bernoulli's law, where when the capillary pipe in a mosquito sprayer is blocked, it causes a zero crossing area so that the liquid in the container cannot go out at all. When the blocked capillary pipe has opened, the pipe has a cross-sectional area so that the mosquito liquid was pumped out well. Students assumed that the liquid can spread out to fill in the room because of the large pressure given by someone who presses the pump rod from the mosquito container causing the speed of pump was large. The difference between the knowledge had by students and the concept of physics taught by their teacher causes misconception.

The misconception experienced by students must be diagnosed early so that such misconception can be stopped. One way to detect misconceptions experienced by students is with a four-tier diagnostic test (Anggrayni and Ermawati, 2019). The four-tier diagnostic test is the most complete diagnostic test with a four-tier arrangement consisting of: 1) the answer option, 2) the answer level of confidence, 3) the choice of the answer reason, and 4) the level of confidence in the reason for the answer. The four-tier diagnostic test can diagnose students misconception in dept by distinguishing the level of conception of students into seven categories based on the combination of answers given by students. The intended categories are given in Table 1.

**Table 1.** Categories of combination four-tier answers (Source: Gurel et al, 2015)

No.	Option	Confidence	Reason	Confidence	Categories
1	Correct	Sure	Correct	Sure	SC
2	Correct	Sure	Wrong	Sure	FP
3	Wrong	Sure	Correct	Sure	FN
4	Wrong	Sure	Wrong	Sure	MSC
5	Correct	Sure	Correct	Not Sure	LKG
6	Correct	Not Sure	Correct	Sure	LKG
7	Correct	Not Sure	Correct	Not Sure	LKG
8	Correct	Sure	Wrong	Not Sure	LKG
9	Correct	Not Sure	Wrong	Sure	LKG
10	Correct	Not Sure	Wrong	Not Sure	LKG
11	Wrong	Sure	Correct	Not Sure	LKG
12	Wrong	Not Sure	Correct	Sure	LKG
13	Wrong	Not Sure	Correct	Not Sure	LKG
14	Wrong	Sure	Wrong	Not Sure	LKD
15	Wrong	Not Sure	Wrong	Sure	LKD

16	Wrong	Not Sure	Wrong	Not Sure	LKD
17	If one, two, three or all are not answered				Cannot be encoded

SC = scientific conception, MSC = misconception, FN = false negatives, FP = false positives, LKG = lack of knowledge guessing, LKD = lack of knowledge deficiency.

Based on the description on the paragraphs above, it is necessary to develop a valid and reliable misconception diagnostic test (Sugiyono, 2015); and that work is reported in this paper. The aim was to determine the validity (both internal and external validities) and the reliability of the developed four-tier diagnostic test instrument for Dynamic Fluid concepts.

## METHOD

A draft of misconception diagnostics test for Dynamic Fluid concepts was written based on the literature study on these concepts and based on the trial test result applied to 40 students (the commencement years of 2017 & 2018) in Physics Department, The State University of Surabaya to collect the most possible answers and reasons on the developed questions. These answers and reasons are than occupied as the 1<sup>st</sup> and 3<sup>rd</sup> tiers of the developed test. The developed four tier test consists at 20 numbers of question was then tested to the intended sample of students i.e. 50 numbers of 11<sup>th</sup> grade students in Science class in Senior High School 1 Sidoarjo to obtain the external validity and the reliability data. Assessment to the internal validity was obtained from the evaluation carried out by two independent Physics Lecturers in Physics Department, The State University of Surabaya. Table 2 provides the range of score of the internal validity and the interpretation.

**Table 2.** Interpretation of Validity Test Results and Criteria (Riduwan and Akdon, 2013: 18)

Score (%)	Validity Interpretation
0 - 20	Very weak
21 - 40	Weak
41 - 60	Enough
61 - 80	Valid
81 - 100	Very valid

Analysis of empirical validity of content is done by calculating the percentage of false positives (FP) and false negatives (FN). FP is combination of the second answer in Table 1. While FN is the third combination of answers in Table 1. According to Hestenes and Halloun (1986), the empirical validity of the content will be fulfilled if the percentage of FP and FN values obtained from the test results less than 10%.

According to Arikunto (2013), empirical validity of constructs for each item and the whole instrument can be calculated using the Pearson product moment correlation

equation ( $r_{xy}$ ). The instrument to be valid empirically if the value of  $r_{xy} > r_{table}$ .

The reliability test is calculated using the Cronbach Alpha equation or  $r_{11}$  (Siregar, 2013). An instrument to be reliable if the value of  $r_{11} > r_{table}$  (Sugiyono, 2015). The  $r_{table}$  value used 0.279 with a sample size of 50 students and a significant level of 5% (Arikunto, 2016).

## RESULTS AND DISCUSSION

Before the diagnostic test instruments were tested in the 50 students, the instrument was validated in advance by two independent Physics Lecturers to find out the score of internal validity. The results of internal validity are shown in Table 3.

**Table 3.** Recapitulation of the results of internal validity

No.	Aspects of Internal Validity	Value (%)
1.	Content	93,75
2.	Construct	95,00
3.	Language	88,89

According to Riduwan and Akdon (2013), the results of the score of the three aspects of the instrument's internal validity that have been developed based on Table 3 are included in very valid criteria with a range of 81% - 100%. This means that the instrument developed is feasible to use.

External validity is obtained by testing instruments that have been validated to 50 students in one of the Science Classes, 11<sup>th</sup> grade in Senior High School 1 Sidoarjo. Data obtained from the results of these trials are then analyzed by the Author to determine the empirical validity of content and empirical validity of the constructs of instruments that have been developed. The results of the calculation of FP and FN for empirical validity of the contents are shown in Table 4.

**Table 4.** Recapitulation of results of FP and FN

Question Number	False Positives (FP)	False Negatives (FN)
1.	-	-
2.	1	-
3.	1	1
4.	1	-
5.	3	3
6.	1	-
7.	2	-
8.	2	1
9.	2	5
10.	5	3
11.	-	1
12.	-	-
13.	7	1
14.	-	2
15.	-	-
16.	8	8
17.	1	1

18.	6	1
19.	1	9
20	2	2
<b>Total (Σ)</b>	43	38
<b>Percentage</b>	4,3 %	3,8 %

Based on the data in Table 4, it can be seen that the score of false positives (FP) and false negatives (FN) obtained from the trial results is less than 10%. These results have fulfilled the criteria for empirical validity of contents according to Hestenes and Halloun (1995) so that it can be concluded that the instruments developed by the Author are stated to be empirically valid contents.

The empirical validity of the construct for each item and the whole instrument. Both of these validities are determined using the correlation of pearson product moment. The results of the empirical validity of constructs for the types of items based on the calculation results are in Table 5.

**Table 5.** Calculation of the value of  $r_{xy}$  for each

Question Number	$r_{xy}$	$r_{table}$	Categories
1	0,205	0,279	Invalid
2	0,490		Valid
3	0,455		Valid
4	0,440		Valid
5	0,312		Valid
6	0,385		Valid
7	0		Invalid
8	0,719		Valid
9	0,638		Valid
10	0,570		Valid
11	0,294		Valid
12	0,359		Valid
13	0,482		Valid
14	0,620		Valid
15	0,443		Valid
16	0,294		Valid
17	0,160		Invalid
18	0,132		Invalid
19	0,368		Valid
20	-0,006		Invalid

Based on Table 5, it can be seen that out of the twenty diagnostic test instrument questions there are five items that are invalid because of the  $r_{xy}$  less than  $r_{table}$ . The five invalid items are in Numbers 1, 7, 17, 18, and 20 so they are not used in calculating reliability.

Empirical validity of the construct of the whole instrument is also calculated using the correlation Pearson product moment equation. Based on the results of calculations using the equation obtained  $r_{xy}$  of 0.591. Thus it can be concluded that the instrument developed was declared empirically valid construct because  $r_{xy} > r_{table}$ .

Instrument reliability ( $r_{11}$ ) is calculated using the Alpha Cronbach formula equation. Based on the

calculation using this formula, the value of  $r_{11}$  is 0.751. This result means that the value of  $r_{11} > r_{table}$  so that it can be said that the instrument developed is reliable.

Based on the analysis above, it can be seen that the four-tier diagnostic test instruments developed have met the criteria of validity and reliability, both internal and external validity. These results are in accordance with the criteria of a good instrument by Sugiyono (2015) so that it can be concluded that the instruments developed are declared valid and reliable.

## CONCLUSION

Based on the results obtained from this study, it can be concluded that the instruments developed were declared valid, both internally and externally. The reliability value ( $r_{11}$ ) obtained is also greater than the  $r_{table}$  value so that it can be stated that the instrument is also reliable. Thus it can be concluded that the four-tier format diagnostic test instrument is feasible to be used in identifying students' misconceptions.

## REFERENCES

- Amin, N., Indonesia, U. P., Samsudin, A., & Indonesia, U. P. (2016). PROSIDING SNIPS 2016 Analisis Instrumen Tes Diagnostik Dynamic-Fluid Conceptual Change Inventory ( DFCCI ) Bentuk Four-Tier Test pada Beberapa SMA di Bandung Raya, (March 2017).
- Anggrayni, S., & Ermawati, F. U. (2019). The validity of Four-Tier ' s misconception diagnostic test for Work and Energy concepts The validity of Four- Tier ' s misconception diagnostic test for Work and Energy concepts. <https://doi.org/10.1088/1742-6596/1171/1/012037>.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: PT. Bumi Aksara.
- Gurel Derya Kaltakci, dkk. (2015). A Review and Comparison of Diagnostic Instruments to Identify Students ' Misconceptions in Science. *11*(5), 989–1008. <https://doi.org/10.12973/eurasia.2015.1369a>
- Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory. *The Physics Teacher*, 502–506.
- Isfara, L., & Ermawati, F. U. (2018). Validitas Instrumen Four-Tier Misconception Diagnostic Test untuk Materi Fluida Statis. *Inovasi Pendidikan Fisika*, 07(03), 429–433.
- Riduwan, & Akdon. (2013). *Rumus dan Data dalam Analisis Statistika*. Bandung: Alfabeta.
- Sugiyono. (2015). *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Suparno, P. (2013). *Miskonsepsi dan Perubahan Konsep dalam Pendidikan Fisika*. Jakarta: PT Grasindo.
- Utari, J. I., & Ermawati, F. U. (2018). Pengembangan Instrumen Tes Diagnostik Miskonsepsi Berformat Four-Tier untuk Materi Suhu, Kalor dan Perpindahannya. *Inovasi Pendidikan Fisika*, 07(03), 434–439.
- Sholihat, F. N., Samsudin, A., & Nugraha, M. G. (2017). Identifikasi Miskonsepsi dan Penyebab Miskonsepsi Siswa Menggunakan *Four-Tier Diagnostic Test* Pada Sub-Materi Fluida Dinamik: Azas Kontinuitas. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3(2), 175–180. <https://doi.org/10.21009/1.03208>