Development of HATRADI: a four tier test for diagnostic student misconception in heat transfer concept

H Y Suhendi and R Ardiansyah

1Program Studi Pendidikan Fisika, UIN Sunan Gunung Djati Bandung, Jl. AH. Nasution No. 105, Bandung 40614, Indonesia
2Jurusan Pendidikan Biologi, Universitas Siliwangi, Jl. Siliwangi No. 24, Tasikmalaya 46115, Indonesia

*Corresponding author: herni.suhendi@uinsgd.ac.id

Abstract. Heat transfer concept is an abstract concept that creates misconceptions. Misconception is an understanding of a concept that is contrary with theory of scientists, if not immediately diagnosed it will cause a prolonged chain of misunderstandings. Diagnostic tests are one solution to identify these misconceptions. This research aims to develop a multiple choice four-tier test instrument for diagnose and identify student misconceptions on heat transfer concept. The research method used is the R&D research method with a 4D research design. The research was conducted at a school in Bandung with a sample of 60 students who were selected by purposive sampling technique. HATRADI is a diagnostic test item to capture students' misconceptions on heat transfer concept. The results of the research stated that HATRADI was a valid and reliable instrument with a very high category. So it can be concluded that HATRADI is a four tier test instrument that is feasible to assess and identify student misconceptions on heat transfer concept.

1. Introduction

Pesman’s research in 2005 states that misconception is an understanding of a concept that is strongly believed but that concept is not in accordance with the scientific concept of experts [1]. Broadly speaking, the causes of misconceptions can be summarized into five groups, namely: students, teachers, textbooks, context and teaching methods [2]. Of these various factors, the preconceptions or initial concepts brought by students are very influential on the occurrence of a misconception. The initial preconception is in the form of a concept or explanation of a phenomenon as they see it with their own eyes in everyday life.

The concept of heat transfer is basically closely related to everyday life, but because the concepts related to heat are abstract or its quantity cannot be observed directly [3], it often causes misconceptions. There have been found various studies related to misconceptions in students of the concept of heat transfer. Researchers found that students often feel confused about the effect of surface color on the rate of radiation heat transfer, for example they understand that the surface of a black object can withstand heat energy so that it emits radiation slower than a white surface [4]. Other researchers noted students’ thoughts that heat transfer by conduction is accompanied by the movement of particles making up the object [5], color is the most important factor in the process of heat transfer by radiation [6]. If these misconceptions are not resolved immediately, it will interfere with students’ understanding in the future. In addition, misconceptions can also hinder the mastery of concepts and cause low student achievement.
so that there needs to be an emphasis on a concept so that misconceptions in students do not occur continuously. Misconceptions that occur not only have an impact on the concept being studied, but can also affect the concepts that will be studied afterwards because the concepts in physics are interrelated with one another. If misconceptions are not corrected as soon as possible, they will continue to higher levels of education and it will be increasingly difficult to overcome them. It is important for teachers to know the misconceptions experienced by students from an early age so that the misconceptions are not sustainable and spread to other concepts [7].

Köse in 2008 stated that there are several ways commonly used to find out students' understanding of concepts and misconceptions, namely in the form of open questions, two-tier tests, concept maps, prediction-observation explanations, interviews about an event or event, interviews about concepts. 8]. So far, the two-tier test is considered more relevant in measuring misconceptions in students compared to interviews, multiple choice, description tests and class discussions [9]. However, based on a literature review on research conducted by Kutluay in 2005, it was found that there were several drawbacks where the use of a two-tier test instrument in detecting misconceptions in students had too high a result for misconceptions because all wrong answers were considered misconceptions [10]. The use of two-tier also has not been able to know the beliefs of students in understanding a concept. In addition, Turker said that even two-tier tests were unable to distinguish between errors due to lack of knowledge (Lack of Knowledge) and errors due to misconceptions, and to distinguish between correct answers based on guessing and correct answers based on original understanding [11].

The limitation of the two-tier test can be overcome significantly by including the level of confidence (Confidence Rating) as an additional level for the two tier test so that a three tier test is developed to detect misconceptions in students. However, the three-tier test is only able to express one confidence rank for answers and reasons, unable to explain whether students have various levels of confidence for the two levels, so a four-tier test is developed to overcome deficiencies in the three-tier test [12]. The advantages of using a four-tier test instrument in detecting misconceptions in students are that it is able to measure differences in students' knowledge, so that they can find out the level of misconceptions experienced by students [13]. Therefore, the researcher tries to make an evaluation tool that is effective and efficient in detecting misconceptions by developing a four-tier test instrument on the concept of heat transfer. This instrument is called HATRADI. HATRADI stands for HeAt TRAnsfer Diagnostic Instrument, in other words HATRADI is an instrument that functions to identify students 'misconceptions about a concept [14] to analyze students' thinking in answering questions even though the answers given are wrong [15] on the concept of heat transfer.

2. Methods
This research method is R&D using a 4-D development design by Thiagarajan et al. in 1974 [16]. However, in this research only development of the instrument. So, the used is 3D which consists of: (1) Define, (2) Design, and (3) Develop. The stages of the research that have been carried out can be seen in the table below:

| 3D Stage | Detail of Stage |
|----------|----------------|
| Define   | Front end analysis, student analysis, task analysis, concept analysis, and formulation of learning objectives that can be used as a basis for the formulation of test questions. |
| Design   | Preparation to design of HATRADI that refer to: 1) indicator of understanding the concept of Anderson's taxonomy; 2) the concept is heat transfer; 3) Four tier test adaptation format from Kaltakci-Gurel et al. and Caleon & Subramaniam |
| Develop  | This stage includes content, construct, empirical validity and reliability testing. Content and construct validity is carried out through reviewing test |
instruments by material experts, evaluation experts, and language experts. Expert validation was carried out by three experts. Furthermore, the questions were tested on a limited scale and analyzed using the Anates version 4.0 application.

3. Results and Discussion

Some of the outputs from the stages in the development of HATRADI in this research can be described as follows:

3.1 Define

Based on the problem identification and data collection that had been done, the researcher decided to develop a diagnostic test instrument for student misconception which was named the HATRADI diagnostic test instrument in the form of a four tier test. The define steps taken by the researcher are as follows: 1) The research subject was carried out on students high school level, so that the answer options were 5 options, 2) Adding fourth tier that contains a confidence rating on the reason for the answer to tier three, 3) There were 20 items for this test instrument, and 4) The concept used is heat transfer which consists of the sub-concepts of conduction, convection, and radiation.

3.2. Design

The design stages in making HATRADI are as follows: 1) Determine the concept, namely heat transfer 2) Make a HATRADI blueprint consisting of general objectives, list the concepts and subconcepts to be tested, make indicators, and determine the number of questions for each concepts and subconcepts, 3) Determine the form of the test, namely four tier test form, dan 4) Write the HATRADI question in the format like Figure 1.

| 1.1 (The question of heat transfer concept) |
|--------------------------------------------|
| A. (answer options)                         |
| B. (answer options)                         |
| C. (answer options)                         |
| D. (answer options)                         |
| E. (answer options)                         |

1.2 Confidence rating for 1.1

1. Sure
2. Not Sure

1.3 The reason for the choice answer in tier 1.1

A. ………………………
B. ………………………
C. ………………………
D. ………………………
E. (Empty options that can be filled in by students themselves)

1.4 Confidence rating for 1.3

1. Sure
2. Not Sure

Figure 1. HATRADI format

3.3. Development

HATRADI was developed from a three tier test instrument made by Pesman in 2005. In this four tier test, a confidence rating is added to the reasons for the answers to tier three. Each item developed consists of four tiers. The first tier is multiple choice questions with one answer key that students must choose. The second tier is the confidence rating for the answers to the first tier. The third tier is the choice of the student's reason for choosing the answer in the first tier. In this third tier, it contains four reasons that have been provided and one reason is open. The fourth tier is the confidence rating for the
reasons chosen by students. This fourth tier is the result of development to find out how confident students are in choosing the reasons for the answer [17]. The number of questions was 20 items with sub concepts, namely conduction, convection, and radiation. The format in HATRADI that has been developed can be seen in Figure 2 below:

Figure 2. Sample of HATRADI item test.

After the questions are complete, the next step is to validate the instrument by the expert. The validation was carried out by three experts. Each question is assessed for its validity with each item consisting of 15 aspects of assessment including content, construction, and language aspects. A detailed assessment of each item is carried out so that the test questions used are truly feasible and can be used to measure what to measure [18], namely students' misconceptions. The developed HATRADI has been declared very feasible by the validator. The results of the expert's assessment are presented in table 2 below:

Table 2. HATRADI Validation Results by Team Experts

| Validation indicator | Assessment criteria                                           | Percentage | Category |
|----------------------|--------------------------------------------------------------|------------|----------|
| Content              | The question based on Anderson’s Taxonomy                    | 100%       | very feasible |
|                      | Suitability of content with expected competency demands      | 75%        | feasible |
|                      | Statements and reasons in question in accordance with the opinion of experts | 100%       | very feasible |
|                      | The picture in question corresponds to the content            | 100%       | very feasible |
|                      | **Average**                                                   | **93.75%** | **very feasible** |
| Construction         | Question items are formulated clearly and firmly              | 100%       | very feasible |
The question item does not provide a clue to the answer key 100% very feasible
Item questions can measure concept understanding 100% feasible
The item questions do not use words that have multiple or uncertain meanings and are not double negative 75% feasible
Question items can diagnose misconceptions 100% very feasible
Homogeneous and logical answer choices 75% feasible
The answer choices do not use an answer that says "all of the above choices are wrong" or "all choices are correct" 100% very feasible

| Average         | 92.85% very feasible |
|-----------------|----------------------|
| Language        |                      |
| Question using language in accordance with the rules of Bahasa | 75% feasible |
| Using communicative language | 75% feasible |
| The question is not using the local language | 100% very feasible |
| The answer choices do not repeat the same word / group of words, unless they constitute one unit | 75% feasible |

| Average         | 81.25% very feasible |
|-----------------|----------------------|

Based on table 2 about the percentage of expert judgment, it can be explained that the results of the material expert's assessment show a percentage of 93.75% which is included in the very feasible or valid category. In the assessment and evaluation expert judgment has a percentage of 92.85% with the category very feasible or valid. Whereas in the assessment of language experts, it has a percentage of 81.25% in the category of feasible or valid. Based on the three expert judgments it can be said that HATRADI is usable or feasible.

Furthermore, testing is carried out externally through limited trials. The results of the empirical validation in this study are used as a reference to see the implementation of the exam using the developed HATRADI. The result of the empirical validity has a value of $r_{xy} = 0.80$. Based on table 3 below, the validity value of 0.80 falls into the very high category. This shows that the variables in the instrument significantly have a very high or very good correlation. A question that are said to be valid or have very high validity are questions that can measure the expected competence [19].

### Table 3. Guilford's (1956) validity category

| Value        | Category                      |
|--------------|-------------------------------|
| $0.80 < r_{xy} \leq 1.00$ | very high validity (very good) |
| $0.60 < r_{xy} \leq 0.80$  | high validity (good)          |
| $0.40 < r_{xy} \leq 0.60$  | moderate validity (sufficient) |
| $0.20 < r_{xy} \leq 0.40$  | low validity (less)           |
| $0.00 < r_{xy} \leq 0.20$  | very low validity (poor)      |
| $r_{xy} \leq 0.00$         | Invalid                       |
Reliability test results have a value of 0.89. Based on table 4 below, the reliability value of 0.89 falls into the very high category, then the results of HATRADI development can be said to have very high stability or consistency, which means that the instrument is reliable and in accordance with the actual conditions. This statement is supported by [20], a reliable instrument means that the instrument used is very good for use in retrieving research data, so that it is able to reveal reliable data.

Table 4. Guilford's (1956) reliability category

| Value    | Category               |
|----------|------------------------|
| 0.80 ≤ r₁₁ ≤ 1.00 | very high reliability |
| 0.60 ≤ r₁₁ ≤ 0.80   | high reliability       |
| 0.40 ≤ r₁₁ ≤ 0.60   | moderate reliability   |
| 0.20 ≤ r₁₁ ≤ 0.40   | low reliability        |
| -0.00 ≤ r₁₁ ≤ 0.20  | very low reliability (not reliable) |

4. Conclusion
This research produces HATRADI, namely the HeAt TRAnsfer Diagnostic Instrument which functions to identify student misconceptions in the form of a four tier test on heat transfer concept. HATRADI fulfills the feasibility indicators of three experts, namely content experts, assessment experts and linguists experts with each having an feasibility percentage of 93.75%, 92.85%, and 81.25%. HATRADI has a very high validity and reliability value with a reliability value of 0.80 and a validity value of 0.89. The results of this study indicate that HATRADI is valid and suitable for use as a diagnostic instrument for misconceptions.

References
[1] Kirbulut Z D and Geban O 2014 Eurasia J. Math. Sci. Technol. Educ. 10 509
[2] Suparno P 2013 Miskonsepsi & perubahan konsep dalam pendidikan fisika Gramedia Widiasarana.
[3] Gurcay D and Gulbas E 2015 Res. Sci. Technol. Educ. 33 197
[4] Prince M, Vigeant M and Nottis K 2016 Chem. Eng. Educ. 5052
[5] Wibowo F C 2017 Asia-Pacific Forum on Science Learning and Teaching 18 1
[6] Suhendi H Y, Mulhayatiah D, Anjani R, Ramdani M A and Ardiansyah R 2018 MS&E 434 12288
[7] Kamilah D S and Suwarna I P 2016 Edusains 8 212
[8] King D and English L D 2016 Int. J. Sci. Educ. 38 2762
[9] Eymur G and Geban Ö 2017 Int. J. Sci. Math. Educ. 15 853
[10] Kanli U 2015 Sci. Educ. Int. 26 148
[11] Wijaya C P, Handayanto S K and Muhardjito M 2016 J. Pendidik. IPA Indones. 5 13
[12] Yan Y K and Subramaniam R 2018 Chem. Educ. Res. Pract. 19 213
[13] Kiray S A and Simsek S 2020 Int. J. Sci. Math. Educ. 20 13
[14] Vishnumolakala V R, Qureshi S S, Treagust D F, Mocerino M, Southam D C and Ojeil J 2018 QScience Connect 2018 1
[15] Alawiayah N Susilaningish E and Sulistyaningsih T 2018 J. Innov. Sci. Educ. 7 122
[16] Ulandari L, Amry Z and Saragih S 2019 Int. Electron. J. Math. Educ. 14 375
[17] Kaltakci-Gurel D, Eryilmaz A and McDermott L C 2017 Res. Sci. Technol. Educ. 35 238
[18] Barak M, Watted A and Haick H 2020 Assess. Eval. High. Educ. 45 212
[19] Rahmani M, Ningsih K and Nurdini A 2015 J. Pendidik. dan Pembelajaran Khatulistiwa 4 2
[20] Martunis M, Khaldun I and Zulfadli 2015 J. Ilm. Mhs. Pendidik. Kim. 4 70