



Rapid Artificial Diagnostic Test in Intelligent Teacher Assistant System (ITAS) to Identify Misconceptions on Kinematics

Sutrisno Sutrisno, Maison Maison, Jefri Marzal, and Wawan Kurniawan^(✉)

Program Studi Doktor Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Jambi,
Jambi, Indonesia

kurniawan_wawan@unja.ac.id

Abstract. Errors in interpreting a concept can occur because students are still in the process of understanding. The experience gained through observation and reasoning has not been able to form complete knowledge, so students often experience errors in interpreting a concept. Therefore, errors in interpreting a concept that is not following the concept of science must be assessed as early as possible by the teacher. This research is a research development that produces a product using a four-level diagnostic test based on artificial intelligence. The purpose of this study was to see the development process and the diagnostic feasibility level of a four-level test using an Intelligent Teacher Assistant System (ITAS) based on the responsive website, the subject of developed Kinematics. This research was conducted at a senior high school in Tanjung Jabung Barat. Four validators of Physics Education lecturers carried out the feasibility test at Jambi University and one physics subject teacher at senior high school in Tanjung Jabung Barat. The result of the feasibility level of the media by media experts is 93.28% (category “very feasible”). In contrast, the test results or user responses obtained a value of 80.43% (category “very feasible”). Based on what was done, the development of a four-level test diagnostic using Matlab, Python Programming, Orange Data Mining, and a responsive website to identify high school student misconceptions on the subject of statistical fluids in total on the aspects of software engineering and visual communication can be categorized as “very feasible”, so that it can be used to support the learning process of students in class.

Keywords: Rapid artificial diagnostic test · Intelligent Teacher Assistant System · Misconceptions · Kinematics

1 Introduction

The body text starts with a standard first-level Before students take part in classroom learning, teachers need to realize that each student already has knowledge from daily experience and information from the surrounding environment, called the student’s initial concept (preconception) [1]. However, the initial concepts that these students have are not all the same as the scientific concepts that have been agreed upon by the experts.

This is usually called a misconception [2]. The existence of misconceptions will hinder the process of acceptance and assimilation of knowledge in students so that it will affect the success of students in the process of further learning. This situation is a big mistake in the physics learning process that cannot be tolerated.

Errors in interpreting a concept that is not following the concept of science (misconceptions) must be known as early as possible by the teacher. This is in line with the function of the diagnostic test, which is to identify problems or difficulties experienced by students so that later they can be used to plan follow-up actions in the form of efforts to solve problems or difficulties that have been identified [3].

One of the diagnostic tests that can identify students' misconceptions is the four-tier diagnostic test [4]. This test consists of the first level for knowledge questions in the form of multiple choice with answer choices, the second level contains the confidence of the answers at the first level, the third level contains the reasons for the answers at the first level, and the fourth level contains the level of confidence, on the reasons for the answers at the third level [5]. The four-tier test can assess misconceptions with a low error rate and cover the lack of knowledge because it can hold all the advantages of the three-tier test [6].

However, diagnostic tests that are easy to use to identify students' conceptual mastery are still not widely developed and used by teachers. The results of research conducted by [7, 8], there are still many physics teachers who have not used and have never measured students' misconceptions, and there are no diagnostic tests available. More specifically, the results of interviews conducted by [9] the teacher only uses daily test tests as an indicator of student competency achievement, and the identification of misconceptions using diagnostic tests is not carried out because the process of correcting student answers is quite time-consuming and the correction process is complicated [10].

Diagnostic tests that are usually prepared are still in the form of paper-pencil tests, so mapping student weakness profiles and providing feedback cannot be done quickly [11]. Another weakness in using the paper-pencil test system is that manual correction is prone to errors, and it is challenging to classify students who understand concepts, misconceptions, and do not understand [12]. The application of this website is expected to make it easier for teachers in preparation, processing, efficient assessment, shortening time, and reducing errors in the assessment process. Besides that, students can also see from each test that has been done per item. This is useful for making it easier for teachers to classify students into several learning teams, and teachers can make decisions for students who have a low conceptual understanding [13].

One alternative that can be done to overcome these problems is to utilize information and communication technology (ICT) which is currently being developed in education, namely by creating a system that can make it easier to make four-tier diagnostic tests on the website. The research results by [14] say that website-based diagnostic tests can be used to diagnose student learning difficulties quickly and accurately.

2 Method

In research on developing the Diagnostic four-tier test using this responsive website, the researcher used the ADDIE model developed by [15–17]. There are five stages in this

Table 1. The score is to use a Likert scale

Score Description	Score
Strongly Disagree	1
Disagree	2
Doubt	3
Agree	4
Strongly Agree	5

Table 2. The Category Validation Sheet

Score Interval (%)	Category Score
0 - 20	Very Inappropriate
21 - 40	Not feasible
41 - 60	Decent enough
61 - 80	Worthy
81 - 100	Very Worthy

development: analysis, design, development, implementation, and evaluation, but in this study, only up to the development stage. The analysis stage begins with conducting a needs analysis with a literature study. At the design stage, it is done by determining the method or planning that will later be used to develop a four-tier diagnostic test using a responsive website. The last stage of development is testing the diagnostic four-tier test using a responsive website developed. The tests carried out are (1) unit testing, (2) integration testing, (3) system testing, and (4) acceptance testing [18].

The types of data used in this research are qualitative and quantitative. Qualitative data is in the form of integration testing and system testing results and validator's suggestions for the diagnostic four-tier test using a responsive website that has been developed. Meanwhile, quantitative data were obtained from the questionnaire scores of media experts by validators and user responses.

The data collection techniques used were black-box testing instruments, installation/launch testing instruments, and media expert instruments. This instrument is an adaptation of [19] adapted to the development carried out. The instrument used has also passed the instrument expert validation stage by two lecturers of Physics Education at the University of Jambi. Furthermore, the data analysis technique used descriptive statistical analysis techniques using a Likert scale. According to [20], one of the most frequently used ways to determine the score is to use a Likert scale. The Likert scale score used is a scale of 5, which is described in Table 1.

The classification of scores from the validation results by media experts is obtained, which is categorized into five ranges, as shown in Table 2.

The data in the inappropriate or very inappropriate category is analysed along with suggestions and comments by media experts to find out the shortcomings of the product. The analysis results are used as material to revise the product, and then the final product will be obtained.

3 Result and Discussion

This development research is to produce a diagnostic four-tier test using a responsive website to identify high school students' misconceptions on the subject of static fluids. This development uses the ADDIE development model, which consists of the stages of analysis (analyze), design (design), development (development), implementation (implementation), and evaluation (evaluate). However, this research was conducted only up to the development stage. This is because this research aims only to develop and produce a suitable product for use.

The first stage in the development of this product is the analysis stage. The analysis stage is the foundation in development to identify the source of the problem and determine the right solution in developing the product to be made [21]. The results of the literature study obtained can be seen as follows: Literature Study Results from No Literature Study Results: (a) There are many misconceptions about fluid mechanics in the sub-concept of hydrostatic pressure and Archimedes' law [22], (b) Students experience misconceptions as much as 39.24% on the concept of hydrostatic pressure, 38.85% on Pascal's law material, and 35.2% on the concept of buoyancy [23], (c) the results of a preliminary study conducted by [2] there are still many physics teachers who have not developed a diagnostic test, (d) teachers never measure students' misconceptions, and no diagnostic tests are available; they only use daily test tests as an indicator of student competency achievement [24], (e) the Four-Tier Diagnostic Test has fulfilled the function of mapping the student's weakness profile; it is just that in this study, the diagnostic test compiled was still in the form of a paper-pencil test (PPT) [24], (f) web-based tests can automatically check the results of working on questions [25], and (g) web-based diagnostic tests can be used to quickly and accurately diagnose student learning difficulties [26].

Based on the description of the results of the literature study above, researchers can design and develop products that will be made and developed in the form of a diagnostic four-tier test using a responsive website to identify high school students' misconceptions on the subject of Kinematics.

Furthermore, the researcher began to design a diagnostic four-tier test using a responsive website at the design stage. The website uses Xampp as a web server and database server, visual studio code as a coding editor for PHP, HTML, and CSS programming languages. The prototype is the initial form of the designed, built, and becomes a raw example of the actual product. Prototypes are made first in this development, which will later be tested to improve concept aspects, engineering functions, product operations, technical, technology, and value aspects. The prototype results are as follows (Figs. 1, 2 and 3).

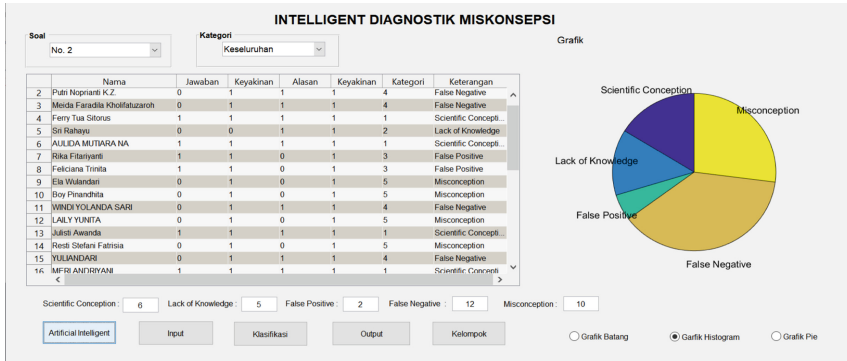


Fig. 1. Development of a misconception measurement system using classification Artificial Intelligence

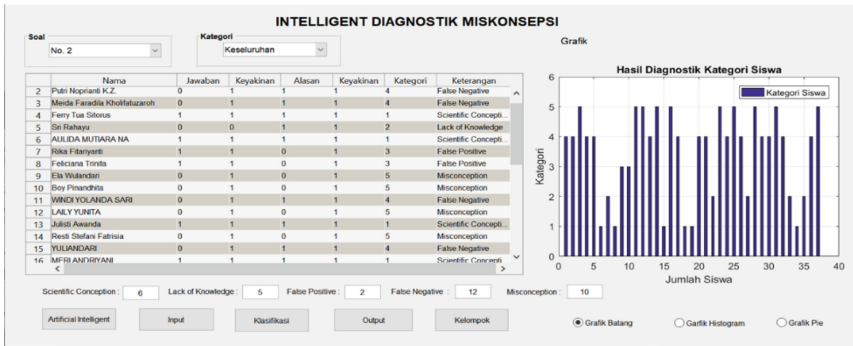


Fig. 2. The results of the diagnosis of high school students' misconceptions

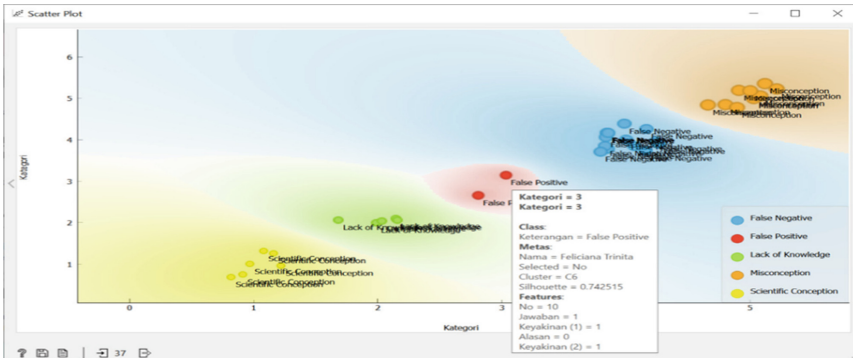


Fig. 3. Misconception's clustering results using the orange platform

Table 3. Media expert validation result

Aspect Assessment	Validation I (%)	Validatio II (%)	Category
Usability	82,50	96,00	Very Eligible
Functionality	76,25	95,50	Very Eligible
Communication Visual	66,55	88,36	Very Eligible
Average	75,10	93,28	Very Eligible

After going through the design stage, the next stage is the development stage. At this stage, the initial draft design of the diagnostic four-tier test using a responsive website that has been produced in the previous stage will be realized in a tangible form, and testing of the products that have been developed will be carried out. The testing is carried out in four stages.

Unit testing results from this test: there are no website links that experience errors and do not experience errors. Integration testing, the result of this test is a test of the program's functionality by using the black-box as a reference for testing (Fig. 3). All tested functionality programs get no errors, and all run fine. After all the programs and website pages have been running well without any errors, system testing will be carried out.

From Table 3, media experts test the quality of the website from the usability, functionality, and visual communication aspects. Validation is carried out twice before entering the trial process or user response. The percentage results obtained in the first validation are 82.50% usability, 76.25% functionality, and 66.55% visual communication. So the acquisition of the average percentage of the first validation results is 75.10%, which is included in the "appropriate" category. After knowing the value and suggestions given in the first validation, the researcher revised the website and then did it again in the second stage of the validation test. The percentages obtained from the second stage of the validation test are 96.00% usability, 95.50% functionality, and 88.36% visual communication. The average percentage of the second validation result is 93.28%. Therefore, the average percentage results obtained from the two validation tests are included in the "very feasible" category.

After passing the media expert validation stage twice, then proceed with the trial phase or user response. At this stage, 120 students of class senior highs schools in Tanjung Jabung Barat were involved. The results of student responses can be seen in Table 4.

Based on the study results obtained in Table 4, the average percentage of each aspect of the user response questionnaire obtained a percentage of 80.85% from the aspect of interest in the website and 80.00% from the aspect of ease of using the website. The average percentage of all aspects is 80.43% which is included in the "very feasible" category.

Based on the media feasibility test consisting of software engineering testing, the results of media expert validation, and trials or user responses, it can be concluded that the diagnostic four-tier test uses a responsive website to identify high school students'

Table 4. User response result analysis

Aspect Assessment	Percentage Eligibility I (%)	Category
100	80,85	Very worth it
150	80,00	Very worth it
Average	80,43	Very worth it

misconceptions on the subject of static fluids that have been developed “very feasible” used to support the learning and learning process in the classroom.

4 Conclusion

The conclusion that can be drawn from the development research that has been completed is that The process of developing a diagnostic four-tier test using a responsive website to identify high school students’ misconceptions on the subject of static fluid consists of the stages of analysis (analyse), design (design) and development (development). Website develop uses Xampp as a web server and database server, visual studio code as a coding editor for PHP, HTML, and CSS programming languages. The product results are in the form of a diagnostic four-tier test using a responsive website. The results of the media feasibility level by media experts are in the “very feasible” category. While the results of trials or user responses, the value obtained are in the “very feasible” category. Based on the tests carried out, the development of a diagnostic four-tier test using a responsive website to identify the misconceptions of high school students on the overall subject of static fluids in the aspects of software engineering and visual communication can be categorized as “very feasible,” so that it can be used to support the learning process of class students. This research still has many shortcomings, and there are still many things that need to be studied and developed again. So that later this website can be used by teachers and students without any errors and can dash, precisely and well.

Acknowledgments. The author would like to thank the assistance or encouragement from colleagues, exceptional work by technical staff, or financial support from Universitas Jambi, The Research and Community Service Institute (LPPM).

References

1. Mufit, F., Festiyed, F., Fauzan, A., Lufri, L.: Impact of learning model based on cognitive conflict toward student’s conceptual understanding. IOP Conf. Ser. Mater. Sci. Eng. 335, p. 012072 (2018). <https://doi.org/10.1088/1757-899X/335/1/012072>.
2. Türkoguz, S.: Investigation of three-tier diagnostic and multiple choice tests on chemistry concepts with response change behaviour. Int. Educ. Stud. (2020), <https://eric.ed.gov/?id=EJ1266489>.

3. Wardani, Y.R., Mundilarto, M., Jumadi, J.: The influence of practicum-based outdoor inquiry model on science process skills in learning physics,” ... Pendidikan Fisika Al core.ac.uk (2019), <https://core.ac.uk/download/pdf/276126023.pdf>.
4. Diani, R., Yuberti, Y., Anggereni, S.: ECIRR (Elicit, Confront, Identify, Resolve, Reinforce) learning method with the pictorial riddle method: is it effective in reducing physics misconceptions?. *J. Phys.*, (2020), <https://iopscience.iop.org/article/https://doi.org/10.1088/1742-6596/1572/1/012020/meta>.
5. Suganda, T., Kusairi, S., Azizah, N.: The correlation of isomorphic, open-ended, and conventional score on the ability to solve kinematics graph questions. *J. Penelit.*, (2020), <http://journal.unj.ac.id/unj/index.php/jpppf/article/view/14905>.
6. Çoban A., Mustafa, E.: Development of three-tier scale insufficiencies of classic physics conceptual comprehension scale (ICPPCS). *Online Science Education Journal* 4(2), (2019), <https://dergipark.org.tr/en/pub/ofed/issue/50430/614730>.
7. Soeharto, Csapó, B., Sarimanah, E., Dewi, F.I., Sabri, T.: A Review of students’ common misconceptions in science and their diagnostic assessment tools,” *J. Pendidik. IPA Indones* 8(2), (2019), <https://doi.org/10.15294/jpii.v8i2.18649>.
8. Widayati, N.T., Wiyanto, W., Subali, B.: Analisis kualitas pengembangan four tier test untuk mengetahui profil kemampuan metakognisi siswa tingkat SMA Eks-Karesidenan Pati. *UPEJ Unnes Phys.*, (2020), <https://journal.unnes.ac.id/sju/index.php/upej/article/view/41926>.
9. Mataka, I., Taibu, R.: A multistep inquiry approach to improve pre-service elementary teachers’ conceptual understanding. *Int. J. Res. Educ.* (2020), <https://eric.ed.gov/?id=EJ1229009>.
10. Ekayati, I.A.S., Efendi, D.I., Sumadi, S.: Pengembangan three tier diagnostic test untuk mengidentifikasi miskonsepsi mahasiswa PAUD. *Pros. SNasPPM* (2020), <https://snasppm.unirow.ac.id/prosiding/index.php/SNasPPM/article/view/315>.
11. Mešić, V., et al.: Measuring students’ conceptual understanding of wave optics: A Rasch modeling approach,” *Phys. Rev. Phys. Educ. Res.* 15(1), pp. 1–20 (2019), <https://doi.org/10.1103/PhysRevPhysEducRes.15.010115>.
12. Afif, N.F., Nugraha, M.G., Samsudin, A.: Developing energy and momentum conceptual survey (EMCS) with four-tier diagnostic test items. *AIP Conf. Proc.*, (2017), <https://aip.scitation.org/doi/abs/https://doi.org/10.1063/1.4983966>.
13. Bozdağ, H.C., Gökçe, O.K.: Dört aşamalı kavramsal ölçme aracı ile sekizinci sınıf öğrencilerinin hücre bölünmeleri konusundaki bilgi farkındalıkları ile kavram yanlışlarının belirlenmesi. *Sak. Univ. J. Educ.* (2018), <https://dergipark.org.tr/en/pub/suje/article/389668>.
14. Liu, M.C., Huang, Y.M.: The use of data science for education: the case of social-emotional learning. *Smart Learn. Environ.* 4(1), p. 1 (2017), <https://doi.org/10.1186/s40561-016-0040-4>.
15. Gurel, D.K., Eryilmaz, A., McDermott, L.C.: A review and comparison of diagnostic instruments to identify students’ misconceptions in science. *Eurasia J. Math. Sci. Technol. Educ.* 11(5) (2015), <https://doi.org/10.12973/eurasia.2015.1369a>.
16. Andiska, G.S.: Remediasi miskonsepsi melalui model SSCS (search, solve, create, share) dengan metode resitasi pada materi suhu dan kalor. *Repository.radenintan.ac.id*, (2018).
17. Gurel, D.K., Eryilmaz, A., McDermott, L.C.: Development and application of a four-tier test to assess pre-service physic teachers’ misconceptions about geometrical optics. *Res. Sci. Technol. Educ.* 35(2), pp. 238–260 (2017), <https://doi.org/10.1080/02635143.2017.1310094>.
18. Akman, Ö.: TPACK survey development study for social sciences teachers and teacher candidates. *Int. J. Res. Educ. Sci.* 1(1), pp. 1–10 (2015), <https://doi.org/10.21890/ijres.97007>.
19. Lei, X.F., Yang, M., Cai, Y.: Educational data mining for decision-making: a framework based on student development theory 117, pp. 628–641 (2017), <https://doi.org/10.2991/eeeis-16.2017.76>.

20. Yukselturk, E., Ozekes, S., Türel, Y.K.: Predicting Dropout Student: An Application of Data Mining Methods in an Online Education Program,” *Eur. J. Open, Distance E-Learning* 17(1), pp. 118–133 (2018), <https://doi.org/10.2478/eurodl-2014-0008>.
21. Koedinger, K.R., D’Mello, S., Mclaughlin, E.A., Pardos, Z.A., Rosé, C.P.: Data mining and education,” *Wiley Interdiscip. Rev. Cogn. Sci.*, 6(4), pp. 333–353 (2015), <https://doi.org/10.1002/wcs.1350>.
22. Diyana, T.N., Sutopo, S.: The effectiveness of web-based recitation program on improving students’ conceptual understanding in fluid mechanics. *J. Pendidik. IPA ...*, (2020), <https://journal.unnes.ac.id/nju/index.php/jpii/article/view/24043>.
23. Cari, C., Pratiwi, S.N., Affandy, H.: Investigation of undergraduate student concept understanding on Hydrostatic Pressure using two-tier test. *J. Phys. ...*, (2020), <https://iopscience.iop.org/article/https://doi.org/10.1088/1742-6596/1511/1/012085/meta>.
24. Aryani, W.D., Suhendi, E., Suyana, I.: Effectiveness of implementation interactive conceptual instruction (ICI) with computer simulation to overcome students’ misconceptions about newton’s law of *J. Phys. ...* (2019), <https://iopscience.iop.org/article/https://doi.org/10.1088/1742-6596/1280/5/052011/meta>.
25. Romero, C., Ventura, S.: Educational data mining and learning analytics: An updated survey,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 10(3), pp. 1–21 (2020), <https://doi.org/10.1002/widm.1355>.
26. Suhirman, X.X., Zain, J.M., Herawan, T.: Data mining for education decision support: A review,” *Int. J. Emerg. Technol. Learn.* 9(6), pp. 4–19 (2014), <https://doi.org/10.3991/ijet.v9i6.3950>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

