ETDC Indonesian Journal of Research and Educational Review
Volume 1, No. 4, 2022, pp 562-570
e–ISSN 2809-3682
p–ISSN 2809-7505
https://doi.org/10.51574/ijrer.v1i4.629

CONSTRUCTING THE FORMATIVE TEST FOR DIAGNOSING MISCONCEPTIONS OF FIRST-SEMESTER HIGH SCHOOL BIOLOGY

Ikhsanudin Ikhsanudin[1], Bambang Subali[2]
[1] Universitas Sultan Ageng Tirtayasa, Indonesia
[2] Universitas Negeri Yogyakarta, Indonesia

Article Info

Article history:

Received August 01, 2022
Revised September 15, 2022
Accepted September 20, 2022

Keywords:

Biology learning
Educational assessment
High school
Misconception
Test construction

ABSTRACT

A misconception is one of the learning obstacles that often occurs in biology class. This study aims to construct a diagnostic instrument to detect misconceptions of biological concepts during the first semester of senior high school and test the quality of the developed instrument. The instrument development model used was a modification of the Oriondo, Dallo-Antonio, and Azwar models. The construction begins with planning the measuring domain and preparing the blueprint for the test instrument. The processes were followed by an instrument seminar, peer review, expert judgment validation, and trying out the test on 885 students. In addition, an unstructured interview with the teachers determines the feasibility of the test instrument. The test content validity according to the Aiken formula is 0.85; the items have different levels of difficulty and good discriminating power; the number of 62 items fits the partial credit model. The reliability of the test is shown by a Kappa coefficient of 0.47. Thus, the test instrument proved to be valid and reliable. Thirteen teachers from eleven schools stated that the test is worthy of use in learning. For the school level, the instrument meets the criteria of a good test instrument.

Corresponding Author:

Ikhsanudin Ikhsanudin,
Faculty of Teacher Training and Education, Universitas Sultan Ageng Tirtayasa.
Email: ikhsanudin@untirta.ac.id

1. INTRODUCTION

Knowledge as one of the learning objectives is a very important aspect for educators to pay attention to. In the knowledge dimension, there is a very popular taxonomy of learning objectives formulated by Bloom and revised by Anderson and Krathwohl (Dettmer, 2005; Näsström, 2009). The taxonomy is also known as the cognitive ability level, starting from the lowest level, which is to remember, understand, apply, analyze, and evaluate, to the highest level, namely creating (Krathwohl, 2002; Krathwohl & Anderson, 2010). In this case, mastery of low cognitive levels underlies mastery of higher cognitive levels. For example, the level of remembering and understanding is a cognitive level that students need to master to be able to apply the knowledge learned. Another example is when students are asked to evaluate; these students need the ability to remember, understand, apply, and analyze related problems in the field of science being studied (Krathwohl & Anderson, 2010; Cullinane & Liston, 2016).

One of the fields of science used in the implementation of learning is biology. The scientific structure of biology contains a lot of conceptual knowledge, which is part of the national education goals and, in a narrow sense, is the purpose of biology learning. Campbell and Reece (2010) describe biological knowledge as knowledge related to living things and the ins and outs of life. Biology includes knowledge of simple to complex things, from concrete to abstract, about living things. Biology is a science with very broad and deep material coverage (Dietz et al., 2019). In formal education, biology is included in the national curriculum and studied from primary to higher education.

The complexity of the material often has the potential to become an obstacle to achieving learning goals, including learning biology. In connection with the goal of mastery of competencies in the form of mastery of factual, conceptual, and procedural knowledge, biology teachers are required to facilitate and guide students in mastering concepts, principles, laws, theories, etc (Nawani et al., 2019). Conceptual knowledge, according to Scott et al. (2011), is a scientific product. In other words, students who study biology need to know the product aspects of science, especially in the scientific field of biology. Because of the large number of science products (biological science), in the learning process of biology, it is often difficult for students to master the whole material.

Köse (2008) reveals that many studies are focused on the problem of the dimensions of knowledge of students in science classrooms. Misconceptions are a common occurrence in science classes. Misconceptions can occur in students at the primary, secondary, and higher education levels and can even occur in teachers or prospective teachers (Reydon, 2021). A misconception is an obstacle or barrier for students to understand scientific symptoms or phenomena, including in the field of biology education (Kumandaş et al., 2019). The results of Boo's research (Boo, 2007) show that misconceptions that occur in a person can last a long time and are difficult to change. Misconceptions persist even with formal science learning. Misunderstandings or misconceptions often occur with abstract concepts such as photosynthesis and respiration (Köse, 2008). Opfer et al. (2012) add that misconceptions are one of the important things to pay attention to in developing concept mastery through learning.

There is a significant relationship between mastery of early biology concepts and subsequent learning. Knowledge of or mastery over early biological concepts is needed to serve as the basis for forming or constructing new knowledge so that learning is more meaningful (Star et al., 2014; Stevens et al., 2021). Opfer et al. (2012) reveal that one of the things to note in developing the ability to master the concept is to avoid misconceptions. In learning biology, Tekkaya (2002) emphasizes that "*misconceptions are a barrier to understanding biology.*" Students' misconceptions about scientific phenomena have been the focus of research from the late 20th century (the 1980s) until now. Kumandaş et al. (2019); Chen et al. (2020) reported that there were student misconceptions in high school biology subjects, both in small schools and in large schools. Furthermore, many overseas researchers focus their research on the misconceptions of students, including in the field of biology.

Based on the research findings described above, it is known that misconceptions have a high potential to occur in biology learning. Bahar (2003) explains that misconceptions are strong enough to persist in a person and are difficult to remove. Therefore, it is important to note, identify, and remediate misconceptions. To identify misconceptions, many researchers use diagnostic tests or diagnostic assessments with specially-made instruments. According to Gurel et al. (2015), various types of diagnostic instruments can be modified to increase the effectiveness of the diagnosis of student misconceptions. Modification of diagnostic instruments can be done in the form of interview instruments, open tests, multilevel multiple-choice tests, and so on. Given the importance of mastery of biological concepts and the importance of misconception diagnosis when learning is still taking place in a formative context, this study aims to develop a biological misconception diagnostic test model that focuses on the new construction of first-semester grade X in the senior high school biology test.

2. METHOD AND DISCUSSION

This research is development research that focuses on the construction of cognitive ability tests. The first phase is an instrument design activity that begins with an analytical study of various theories regarding the steps for the preparation of a diagnosis assessment instrument for misconceptions as well as a theory about the alignment of diagnostic tests with formative assessment, then continues with the initiation of initial instrument preparation. The second phase is the construction stage of the misconception diagnosis instrument in high school biology subjects using the steps formulated in the first phase. The development model uses a modification of the Oriondo and Dallo-Antonio test development models (1998) and the construction steps of the Azwar test (2016), which are limited to the test construction. The trial activity was carried out on 885 students of class X from eleven high schools who were drawn from a hypothetical population in the province of Yogyakarta Special Region using a selective sampling technique. The sample of learners from this hypothetical population is convenient or easy to find in the field as subjects that can be diagnosed. To investigate the validity of the instrument content, data in the form of expert assessment scores were collected through expert assessment sheets and analyzed using the Aiken formula. Furthermore, to test the quality of the instrument, data in the form of students' responses to the test were obtained from the implementation of test trials. The reliability of the test is determined by the Kappa coefficient. The empirical validity was determined using the fitness item (infit mean-square / infit MNSQ). Item characteristics were analyzed using the partial credit model and classical test theory. Finally, the teacher's responses to the instruments compiled were obtained through unstructured interviews after test trials to support the instrument's feasibility.

3. RESULTS AND DISCUSSION

3.1. Description of Constructed Test

The products compiled in this study were a formative test instrument for high school biology grade X on the first semester in the form of two test packages, each consisting of 35-item questions with 7 item items as anchors (joint items). The instrument in question is intended to diagnose learning difficulties in the form of students' misconceptions about the biological material that has been studied at that level. This misconception can be detected through the analysis of the results of measuring students' understanding of the related material. The scope of the test material based on the curriculum includes (1) Biology as A Science, (2) Biodiversity, (3) Classification of Living Things, (4) Viruses, (5) Bacteria, (6) Protists, and (7) Fungi. In a formative context, detected learning difficulties are expected to be overcome through ongoing learning activities. Indicators and test grids have been written at the level of understanding of biology material in the first semester of senior high school grade X at the beginning of the preparation of the instrument.

The initial instrument was then discussed in a seminar to get input or suggestions for improvement. Based on the results of the instrument seminar, improvements are focused on grammar and scoring guidelines related to answer keys. After the instrument seminar, a peer review was conducted with one Biology Education master program student and one Educational Research and Evaluation master program student with a Bachelor of Biology Education background to improve substantive aspects and the fitness of answer keys. Furthermore, qualitative expert assessments are carried out related to the aspects of substance, construction, and language as well as quantitatively to validate the contents of the instrument. One of the most important things in developing an instrument is proofing validity (El Islami et al., 2019; Ikhsanudin & Subali, 2018). Based on the qualitative advice from the experts and the results of the content validity analysis, nine items were corrected, namely 2, 6, 7, 10, 11, 25, 26, 29, and 33 for both packages A and test package B. In general, at the beginning of the preparation of the instrument, the content validity of the A test package was 0.82 and the B test was 0.83 according to the Aiken validity index.

Two test packages for instrument repair results based on the results of expert assessments were then tried out on 885 high school class X students from eleven schools in Gunungkidul Regency, Yogyakarta Special Region. The test trial is conducted in collaboration with the principal and the biology teacher who teaches the class concerned. Test trials are carried out under the lesson schedule so that they do not interfere with the natural conditions of students in the field. The time for testing the test in each class is two hours of instruction (90 minutes) according to the allocation.

3.2. Test Reliability Estimation

The general reliability of the test from the analysis output using the Quest program is 0.84. These results can be said to be the same as the output analysis using SPSS above. Furthermore, this reliability quantity is used to estimate the Kappa coefficient with the help of the Subkoviak estimation table (Subali, 2016). Based on the reliability value (Cronbach's alpha coefficient) obtained and taking into account the standard score (z-score), the estimated Kappa coefficient is 0.47. According to Subkoviak, the estimated Kappa coefficient is in the satisfy category (Subali, 2016). Thus, the test instruments arranged can provide reliable measurement results.

Table 1. Result of Estimation of Cronbach Reliability

Reliability Statistics of Test A		**Reliability Statistics of Test B**	
Cronbach's Alpha	N of Items	Cronbach's Alpha	N of Items
0.836	35	**0.836**	35

3.3. Empirical Validity of Test (Item Fitness)

The output of the item analysis using the Quest program provides information on the validity of the items in the tests carried out in the form of fitness items according to the MNSQ infit. Item fit is indicated by the MNSQ infit amount ranging from 0.77 to 1.30 (Maley & Bond, 2011; Hidayat & Fadillah, 2019). Based on this range, there is only one item that is not fit, namely item number 7 in the test package B. Overall, because almost all items are in the fit category, it can be said that the test instruments arranged can provide a valid score.

3.4. Items Characteristics of Tests

Table 2. Category of Item Difficulty Index Based on Partial Credit Model (PCM)

Category*	Item Number	N	%
> 2; Very Difficult (*Not Good*)	-	-	-
(-2) to 2 ; Medium (*Good*)	1 to 63 (All items)	63	100
< (-2); Very Easy (*Not Good*)	-	-	-
	Total	63	100

Sources: (Aristiawan et al., 2018; Mardapi, 2017)

Table 3. Category of Item Difficulty Index Based on Classical Test Theory

Category*	Item Number	N	%
< 0.30 ; Difficult	5, 9, 13, 22, 27, 33, 48, 50, 53, 54, 62	11	17.46
0.30-0.70; Medium	2, 6, 8, 10, 11, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 28, 29, 30, 31, 32, 34, 35, 36, 37, 41, 42, 45, 46, 49, 51, 52, 55, 56, 57, 58, 59, 60, 61, 63	41	65.08
>0.70 ; Easy	1, 3, 4, 7, 12, 38, 39, 40, 43, 44, 47	11	17.46
	Total	63	100.00

Source: (Subali, 2016; Mardapi, 2017)

Table 4. Category of Item Discrimination Index Based on Classical Test Theory

Category*	Item Number	N	%
0.71 - 1.00; Very Good	-	-	-
0.41 - 0.70; Good	2, 3, 10, 11, 13, 14, 15, 23, 24, 26, 28, 29, 30, 36, 40, 43, 46, 47, 49, 52, 55, 56, 57, 58, 59, 63	26	41.27
0.20 - 0.40; Fair	1, 4, 5, 6, 7, 8, 9, 12, 16, 17, 18, 19, 20, 21, 22, 27, 31, 32, 33, 34, 35, 37, 38, 39, 42, 44, 45, 48, 50, 51, 53, 54, 60, 61, 62	35	55.56
0.00 - 0.20; Bad	25, 41	2	3.17
< 0 ; Very Bad	-	-	-
	Total	63	100.00

Source: (Subali, 2016)

Volume 1, No 4, 2022, pp 562-570

Overall, the test items can be completed by a sample of test subjects within a predetermined time allocation for the test so there is no need to revise the number of test items for that time allocation. From the analysis of the results of the instrument trial, it is known that there is only one item that is not fit, namely item number 7 in test B with an infit amount of 1.40 MNSQ. According to Sumintono and Widhiarso (2015), this test item can be maintained because it does not reduce the quality of the instrument and can still be classified in the good category. With this explanation and with the consideration that most of the test items fall into the fit category, one item that is not fit is corrected and can be included in the measurement using tests that have been completed. The results of the complete improvement of the test construction in this study are printed separately from this article.

4. CONCLUSION

Based on the results of the arrangement of the instruments and data analysis in the research conducted, it can be concluded that the completed instrument is a formative test instrument in an open-reasoned multiple-choice format consisting of two test packages, each consisting of 35 item questions with seven joint items (anchors), and can be used for the diagnosis of misconceptions of biology material in senior high school on the first semester of grade X. According to the Aiken index scale calculated from the assessments of seven experts, the content validity of the two test packages was high, namely 0.85. The empirical validity of the test instrument is shown by the fitness item of the instrument test results. There is only one item that does not fit, but it does not reduce the quality of the instrument. The reliability of the test, according to the Kappa coefficient, is 0.47. The level of difficulty according to item response theory (IRT) proves that the level of difficulty of all items is categorized as "good" with varying levels.

The classical test theory provides information that the level of difficulty of the items is easy (11 items), moderate (41 items), and difficult (11 items), and the difference between all items is positive. All the responding biology teachers stated that the instrument was suitable for use in formative tests to detect student misconceptions. This proves that the test instrument is of good quality. Biology teachers can carry out measurements using instruments that have been prepared as a first step to diagnose students' misconceptions about biology material in the first semester, grade X, in the context of formative assessment. Teachers can also arrange their question items based on the indicators listed on the instrument grid and independently try them out on their students. Biology teachers and/or researchers can analyze test items more deeply by paying attention to the characteristics of the sample of test takers; for example, based on gender, item analysis of male and female students can be carried out, which is not strictly controlled in this study. Analysis of test results can also use other models, such as the two- or three-parameter item response theory model.

REFERENCES

Aristiawan, A., Retnawati, H., & Istiyono, E. (2019). Analysis of Model fit and Item Parameter of Mathematics National Examination Using Item Response Theory. *JPP (Jurnal Pendidikan dan Pembelajaran)*, *25*(2), 40-46. http://dx.doi.org/10.17977/um047v25i12018p040

Bahar, M. (2003). Misconceptions in biology education and conceptual change strategies. *Educational Sciences: Theory & Practice*, *3*(1), 55-64.

Azwar, S. (2016). *Construction of cognitive ability tests (first edition)*. Yogyakarta: Pustaka Pelajar.

Boo, H.K. (2007). Primary science assessment item setters' misconceptions concerning biological science concepts. *Asia-Pacific Forum on Science Learning and Teaching, 8*(1), 1-13.

Campbell, N.A., & Reece, J. B. (2010). *Biology Volume 1 (eighth ed.)*. Jakarta: Erlangga (The original edition was published in 2008 by Pearson Education, Inc.).

Carin, A.A., & Sund, R. B. (1989). *Teaching science through discovery*. Columbus: Merrill Publishing Company.

Chen, C., Sonnert, G., Sadler, P. M., & Sunbury, S. (2020). The impact of high school life science teachers' subject matter knowledge and knowledge of student misconceptions on students' learning. *CBE—Life Sciences Education*, *19*(1), ar9. https://doi.org/10.1187/cbe.19-08-0164

Cullinane, A., & Liston, M. (2016). Review of the Leaving Certificate biology examination papers (1999–2008) using Bloom's taxonomy–an investigation of the cognitive demands of the examination. *Irish Educational Studies*, *35*(3), 249-267. https://doi.org/10.1080/03323315.2016.1192480

Dettmer, P. (2005). New blooms in established fields: Four domains of learning and doing. *Roeper review*, *28*(2), 70-78. https://doi.org/10.1080/02783190609554341

Dietz, R., Letcher, R. J., Desforges, J. P., Eulaers, I., Sonne, C., Wilson, S., ... & Vikingsson, G. (2019). Current state of knowledge on biological effects from contaminants on arctic wildlife and fish. *Science of the Total Environment*, *696*, 133792. https://doi.org/10.1016/j.scitotenv.2019.133792

El Islami, R. A. Z., Sari, I. J., Sjaifuddin, S., Nurtanto, M., Ramli, M., & Siregar, A. (2019, February). An assessment of pre-service biology teachers on student worksheets based on scientific literacy. In *Journal of Physics: Conference Series* (Vol. 1155, No. 1, p. 012068). IOP Publishing. https://doi.org/10.1088/1742-6596/1155/1/012068/meta

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(5), 989-1008. https://doi.org/10.12973/eurasia.2015.1369a

Hidayat, S., & Fadillah, E. N. (2019). Development of assessment instruments in measuring critical thinking skills of senior high school participants of biology subject. In *Journal of Physics: Conference Series* (Vol. 1241, No. 1, p. 012040). IOP Publishing.

Ikhsanudin & Subali, B. (2018). Content validity analysis of first semester formative test on biology subject for senior high school. *Journal of Physics: Conf. Series*, 1097 (2018) 012039. https://doi.org/10.1088/1742-6596/1097/1/012039/meta

Köse, S. (2008). Diagnosing student misconceptions: Using drawings as a research method. *World Applied Sciences Journal*, *3*(2), 283-293.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, *41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Krathwohl, D. R., & Anderson, L. W. (2010). Merlin C. Wittrock and the revision of Bloom's taxonomy. *Educational psychologist*, *45*(1), 64-65. https://doi.org/10.1080/00461520903433562

Kumandaş, B., Ateskan, A., & Lane, J. (2019). Misconceptions in biology: a meta-synthesis study of research, 2000–2014. *Journal of Biological Education*, *53*(4), 350-364. https://doi.org/10.1080/00219266.2018.1490798

Maley, C., & Bond, T. G. (2011). Measuring cognitive development in early childhood learning environments. In *Applications of Rasch Measurement in Learning Environments Research* (pp. 301-324). Brill.

Mardapi, D. (2017). *Educational measurement, assessment, and evaluation (second edition)*. Yogyakarta: Parama Publishing.

Nawani, J., von Kotzebue, L., Spangler, M., & Neuhaus, B. J. (2018). Engaging students in constructing scientific explanations in biology classrooms: a lesson-design model. *Journal of Biological Education*. https://doi.org/10.1080/00219266.2018.1472131

Näsström, G. (2009). Interpretation of standards with Bloom's revised taxonomy: A comparison of teachers and assessment experts. *International Journal of Research & Method in Education*, *32*(1), 39-51. https://doi.org/10.1080/17437270902749262

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, *49*(6), 744-777. https://doi.org/10.1002/tea.21028

Oriondo, L.L. & Dallo-Antonio. (1998). *Evaluating educational outcomes (test, measurement, and evaluation) (5th ed.)*. Quezon City: REX Printing Company.

Reydon, T. A. (2021). Misconceptions, conceptual pluralism, and conceptual toolkits: bringing the philosophy of science to the teaching of evolution. *European Journal for Philosophy of Science*, *11*(2), 1-23. https://doi.org/10.1007/s13194-021-00363-8

Scott, P., Mortimer, E., & Ametller, J. (2011). Pedagogical link-making: a fundamental aspect of teaching and learning scientific conceptual knowledge. *Studies in Science Education*, *47*(1), 3-36. https://doi.org/10.1080/03057267.2011.549619

Stevens, L., Kopnina, H., Mulder, K., & De Vries, M. (2021). Biomimicry design thinking education: A base-line exercise in preconceptions of biological analogies. *International Journal of Technology and Design Education*, *31*(4), 797-814. https://doi.org/10.1007/s10798-020-09574-1

Subali, B. (2016). *Test development: along with empirical validity and reliability investigations*. Yogyakarta: UNY Press.

Sumintono, B. & Widhiarso, W. (2015). *Rasch modeling application on educational assessment*. Cimahi: Trim Komunikata

Tekkaya, C. (2002). *Misconceptions as barrier to understanding biology*. Ankara: Hacettepe Universitesi Egitim Fakultesi Dergisi.