Development of web-based three tier multiple choice test to measure student’s tree thinking; try out

S Julaeha¹*, T Hidayat², and N Y Rustaman²

¹Program Magister Pendidikan Biologi, Universitas Pendidikan Indonesia, Jl. Dr. Setia Budhi No. 229, Bandung 40154, Indonesia
²Departemen Pendidikan Biologi, Universitas Pendidikan Indonesia, Jl. Dr. Setia Budhi No. 229, Bandung 40154, Indonesia

*Corresponding author’s e-mail: siti.julaeha2912@upi.edu

Abstract. Development of three tier multiple choice tests were conducted to detect and measure of student’s tree thinking on classification concept. Three tier multiple choice test was developed with web integration of a Google Classroom application for easy feedback. The study involved a number of high school students (n = 29) who had studied the concept of Animal Classification. The research procedure was carried out through stages: literature study, problem analysis, and instrument design. The instrument was validated by expert and analyze using ANATES version 4.0. Tree thinking was measured on the basis of identification of evolutionary characters, phylogenetic relationships, concept of clade, MRCA (Most Recent Common Ancestor), sister groups, tree topologies, number of evolutionary changes, and sequence of evolution in species. The validation results show that the three tier multiple choice test can be used to measure students’ tree thinking. In addition, the results analysis of items with ANATES version 4.0 indicated that the test has high reliability (r = 0.69), very good of differences (D> 0.4) on 12 items, and difficult on 14 items. This study show that 11 test items were directly accepted (60%), 3 items were accepted but with revision (15%), 4 items were rejected (25%). The conclusion obtained from this study was that the web-based three-tier multiple choice test developed was good and can be used to measure of student’s tree thinking.

1. Introduction
Classification of living things is one of the biological concepts studied in 10th grade in senior high school. In this concept, students are expected to have the ability to apply the principles of phenetic and phylogenetic classification in the Monera, Protista, Fungi, Plantae, and Animalia. Meanwhile, there are still many students who cannot understand the principle of classification, especially in making phylogenetic trees. The phylogenetic tree, also known as the cladogram, phylogeny, and genetic tree is a branched diagram that represents the evolutionary history of a group of organisms [1]. The phylogenetic tree can also be interpreted as a visual representation that describes an evolutionary relationship between groups nested in a taxon [2]. In general, students find it difficult to place taxon in phylogenetic trees. This results in the phylogenetic tree made by students show that they were unable to describe the actual evolutionary relationship in each taxon [3]. Besides, students also have misconceptions when interpreting the relevance of species based on the closeness of the branch end [4].
A very important skill possessed by students to study the phylogenetic tree is called "tree thinking" [5]. Tree thinking is defined by [6] as a way to understand taxon not as an independent replication in a class but as interconnected parts of the evolutionary tree. Students can not only understand phylogenetic trees, they should also be able to construct or construct a phylogenetic tree well [7]. To construct or make a phylogenetic tree they should focus on the key characters possessed by some organisms because these characters show the lineage of a common ancestor [8].

To measure the ability of tree thinking, a valid and reliable test is needed. The existence of the test obtained information on the level of ability of the students’ tree thinking, so that the teacher can plan the right strategy or method for the success of the learning process. In general, the test used to measure tree thinking is an open ended / essay question because it can show all indicators of ability to be assessed. However, the use of multiple choice tests can also be used as an alternative to measure this ability in the presence of additional reasons for the answers chosen by students [9]. The development of this reasoned multiple choice test is due to several disadvantages of multiple choice tests, namely allowing students to provide answers that are not in accordance with their abilities or in other words the answers are in the form of guessing answers of students [9]. Besides, answers that tend to guess at multiple choice tests can affect the reliability of the test [10]. Therefore, a web-based three tier multiple choice test was developed as an alternative test to measure the ability of students’ tree thinking. Three tier multiple choice tests were developed with web integration in the form of a Google Classroom application for easy feedback. The existence of three tier multiple choice tests can provide further explanation and triangulation of student answers, so that the teacher can identify the misconceptions and understanding of students’ concepts [11].

2. Methods
This study used the descriptive method with aimed of describing results of developing a web-based three tier multiple choice test that can measure student’s tree thinking. The subjects used were a number of high school students in 10th grade (n = 29) who had studied the concept of Animal Classification. The study was conducted in one of the high schools in Kabupaten Bandung Barat on April 2019. The procedure of the study was carried out through stages: literature study, problem analysis and instrument design. The instrument was then validated by the expert judgement and analyzed with the ANATES version 4.0. Analysis using the ANATES version 4.0 was carried out after administering the test to students. The purpose of validation and analysis of test items was to obtain good test items and could be used to measure of student’s tree thinking.

3. Result and Discussion
3.1. Design of Web-based Three Tier Multiple Choice Test and Test Item Validation
A web-based three tier multiple choice test was a multiple-choice test that equipped with the option reason answer and the confidence level consists of four options. The number of test items was 20 which contain indicators of the tree thinking to be measured. The measured ability indicators of tree thinking include: The ability of tree thinking measured includes identification of evolutionary characters (i), phylogenetic relationships (ii), concept of clade (iii), MRCA (Most Recent Common Ancestor) (iv), sister groups (v), tree topologies (vi), number of evolutionary changes (vii), and sequence of evolution in species (viii) [12]. All test items were created in Google Form, then distributed to students through the Google Classroom application. Using this application so that feedback in the learning process will be more easily accepted. The existence of feedback given both directly and in writing will help students to reduce errors that occur during the learning process so that student’s abilities can increase or develop [13]. Design of web-based three tier multiple choice test can be seen in Figure 1.
Figure 1. Shows the design of one type of the question contained in a web-based three tier multiple choice test, this type of question was equipped with a phylogenetic tree image as a reference for several questions. The phylogenetic tree image was also equipped with a description of letters and numbers, each of which shows the species and character of evolution. The other type of question can be a table of progress of evolutionary characters in the species used to reference several questions. In this type of problem, students were asked to made a reconstruction of the phylogenetic tree and some other indicators of the tree thinking that allow students to ask. Each image and table were used as references for 2-3 questions. For example, in Figure 1. There was a phylogenetic tree image format equipped with information in the form of characters and species. The phylogenetic tree image was used for 3 questions, namely questions about the character of evolution (question 1), questions about phylogenetic relationships (question 2), questions about MRCA (Most Recent Common Ancestor) (question 3).

Design of the test was then validated by the supervisor and expert. The validation results showed that the web-based three tier multiple choice test can be used to measure student’s tree thinking, with improvements in the form of explanations of biological terms that were considered difficult for students to understand. This was consistent with the statement that knowledge on animal concepts, especially on species identification skills, was based on memorization in the form of biological terms that were considered difficult [14].

3.2. Analysis Test Item using the ANATES Version 4.0
Analysis of test items was carried out after trying out the test to a number of students (n=29) who had studied the concept of animal classification. In implementing the test, students' answer data were obtained which will then be analyzed using the ANATES Version 4.0. Student’s answer data shows eight variation categories of answers, namely as follows: BBY, BBT, BSY, SBY, SSY, BST, SBT, and SST. BBY category, shows correct student answers, wrong reasons, and sure. BBT category, show correct answers, correct reasons, not sure. BSY category, showing correct answers, wrong reasons, and sure. SBY category, shows wrong answers, right reasons, and sure. SSY category, showing wrong answers, wrong reasons, and sure. BST category, showing correct answers, wrong reasons, unsure. SBT category, showing wrong answers, correct reasons, unsure. SST category, showing wrong answers, wrong reasons, unsure. Variation categories in student’s answers can affect the ability score of tree thinking obtained based on scoring guidelines on this issue. Student’s answer data was then input to the ANATES Version 4.0 to find out the quality of the test items. The results of the analysis of test items in the application can be seen in Figure 2.
Figure 2. Diagram of Analysis Test Items Using the ANATES Version 4.0

Figure 2. Shows the percentage of test items accepted based on the results of the analysis using ANATES Version 4.0. The test items accepted were based on validity analysis, 10 test items were accepted and declared significant (50%), 14 test items were accepted as difficult that was in the category of easy, moderate, and difficult (70%), based on differences analysis with a very good (65%). A Good test items should not be too difficult and not too easy. That was in the range of easy, moderate, and difficult [15]. The range of discrimination index criteria is 0.2-0.6 which is moderate, good enough, and very good. For the test items that have a range of criteria below 0.2 must be replaced or revised [16]. Besides, the reliability analysis of test items show a good value of 0.69. Based on the results of the analysis, the recapitulation of test items accepted from the 20 test items developed can be seen in Figure 3.

Figure 3. Recapitulation of Number of Test Items Directly Accepted, Revision, and Rejected

Figure 3. shows that the quality of test items from 20 items developed based on the analysis of validity, reliability, difficult, and differences were 11 test items were directly accepted (60%), 3 items were accepted but with revision (15%), 4 items were rejected (25%). Items of directly accepted as a test were validity, discrimination and difficulty index in good criteria. Items revision that there is one of component is not good criteria, while the rejected items were the three components including validity, discrimination and difficulty index in bad criteria [17]. These results showed the test developed could be used as a tools to measure student's tree thinking. Previous research has developed the same test on the concept of ecology [18] and acid-base concept [19] which shows that the three tier multiple choice test was very good for applying in learning. Because there were tier-2 and tier-3 which could be detail describe of the student's abilities than a two tier multiple choice that does not describe the level of student confidence in the answer [20]. So, this test could be describe the student's prior knowledge and final abilities after learning.
4. Conclusion
The conclusion obtained from this study was that the web-based three-tier multiple choice test developed was good and can be used to measure of student’s tree thinking, but a little improvement was needed on the Animalia concept in the form of foreign terms used in test items. Improvements were also made on some test items that were considered to not meet the reliability and validity of the test.

5. References
[1] Campbell N A, Reece J B, Urry L A, Cain M L, Wasserman S A, Minorsky P V and Jackson R B 2012 Biologi Edisi Kedelapan Jilid 2 (Jakarta: Erlangga)
[2] Dees J and Momsen J L 2016 Student construction of phylogenetic trees in an introductory biology course Evol. Educ. Outreach 9, 1
[3] Sa’adah S, Tapilouw F S and Hidayat T 2017 Undergraduate Students’ Difficulties in Reading and Constructing Phylogenetic Tree J. Physics Conf. Series 812(2017)
[4] Eddy S L Crowe A J Wenderoth M P and Freeman S 2013 How should we teach tree-thinking? An experimental test of two hypotheses Evol. Educ. Outreach 6, 13
[5] Davenport K Milks K J and Tassell R Van, 2015 Investigating Tree Thinking & Ancestry with Cladograms Am. Biol. Teach. 77, 3 p. 198–204
[6] Halverson K L 2011 Improving Tree-Thinking One Learnable Skill at a Time Evol. Educ. Outreach 4, 1 p. 95–106
[7] Johnson G B 2003 The Living World (St. Louis, Missouri: The McGraw-Hill Companies).
[8] Sa’adah S 2017 Perkuliahan Zoologi Vertebrata Berbasis Representasi Filogenetik dan Team-Based Learning untuk Membekalkan Keterampilan Berpikir Kladistik, Berpikir Kritis, dan Penguasaan Konsep (Bandung: Sekolah Pascasarjana Universitas Pendidikan Indonesia)
[9] Sagap, Husain S N and Djirimu M 2014 The Analysis of Conceptual Understanding in Biology Subject by Using Multiple Reasoned Choice on Cell Topic to The Student of Class XI IPA at SMAN 1 Dampal Selatan J. E-Jipbiol 2, 3 p. 1–8
[10] Klufa J 2015 Multiple Choice Question Tests–Advantages and Disadvantages http://www.inase.org
[11] Pacala F A 2018 Development and Validation of Three-Tier Multiple Choice Test for Conceptual Understanding in Momentum and Collision J. Multidisciplinary Approach and Studies 5, 2
[12] Novick L R and Catley K M 2016 Fostering 21st-century evolutionary reasoning: Teaching tree thinking to introductory biology students CBE Life Sci. Educ. 15, 4 p. 1–12.
[13] Abbott J et al 2008 Technology-Based Assessments Improve Teaching and Learning www.setda.org
[14] Jeno L M Grytnes J A and Vandvik V 2017 The effect of a mobile-application tool on biology students’ motivation and achievement in species identification: A Self-Determination Theory perspective Comput. Educ. 107 p. 1–12.
[15] Arifin Z. 2011 Evaluasi Pembelajaran (Bandung: Remaja Rosdakarya, Bandung).
[16] Endrayanto H Y S and Harumurti Y W 2014 Penilaian Belajar Siswa di Sekolah (Daerah Istimewa Yogyakarta: Kanasius)
[17] Zainul A, 2002 Penilaian Hasil Belajar http://staffnew.uny.ac.id
[18] Liampa V George N and Malandrakis 2019 Development and Evaluation of a Three-Tier Diagnostic Test to Assess Undergraduate Primary Teachers’ Understanding of Ecological Footprint Springer 49 pp 711-736
[19] Cetin-Dindar A and Geban O, 2011 Development of a three-tier test to assess high school students’ understanding of acids and bases Proc. 15 pp 600-604
[20] Widivatmoko A and Shimizu K 2018 The Development of Two-Tier Multiple Choice Test to Assess Students’ Conceptual Understanding about Light and Optical Instruments J Pendidikan IPA Indonesia (JPII) 7, 4 p. 491-501
Acknowledgments
Authors wishing to thanks to the school principal and the biology teacher at the research location who have provided research opportunities and permission.