

Rasch Model Assessment for Bloom Digital Taxonomy Applications

Mohd Effendi Ewan Mohd Matore*

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, 43600, Malaysia

*Corresponding Author: Mohd Effendi Ewan Mohd Matore. Email: effendi@ukm.edu.my

Received: 25 December 2020; Accepted: 12 February 2021

Abstract: Assessment using Bloom's taxonomy levels has evolved in a variety of contexts and uses. In the era of the COVID-19 pandemic, which necessitates use of online assessment, the need for teachers to use digital-based taxonomy skills or Bloom's Digital Taxonomy (BDT) has increased even more. However, the existing studies on validity and reliability of BDT items are limited. To overcome this limitation, this study aims to test whether BDT has good psychometric characteristics as a teacher's self-assessment tool using the Rasch model analysis and to investigate the pattern of BDT usage in teaching and learning. By using a quantitative online survey design, this study involves six levels of BDT, namely, Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating. The questionnaire was developed and validated by two experts prior to administration. A stratified random sampling technique was conducted on 774 secondary teachers from five geographical zones in Malaysia, and the Rasch model was analyzed using WINSTEPS 3.71 software. The performances of items improved by Rasch psychometric assessment including the application of BDT among teachers. The hierarchy level was also assessed through graphical analysis, including the Wright map and bubble chart, to demonstrate the powerful performance of the Rasch model analysis in investigating item quality and reliability. Overall, these empirically validated items using the Rasch model could advance the academic knowledge of BDT for future assessment and promote the Rasch calibration in an educational setting.

Keywords: Bloom's digital taxonomy; assessment; dichotomous score; Rasch

1 Introduction

The term "taxonomy" originates from the Greek words "taxis" and "nomos," which refer to "order" and "method," respectively. This term may be referred to as an arrangement or a law in a specific order that is borrowed from biology, which allows certain classifications of the order. In the development of effective methods to perform mental operations, the notion of ordering is essential to classify these operations and skills and to determine the formation sequences in order to grow and solve certain problems [1]. Bloom's taxonomy was introduced



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

in 1965 by Benjamin Bloom [2] based on his idea of pedagogical taxonomy that refers to the development of a clear pedagogical goal system that focuses on cognitively, affectivity and psychomotor development. There are six cognitive activity levels surrounding the cognitive domain of Benjamin Bloom's taxonomy, which range from the easiest to the hardest levels, including Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. The taxonomy was prudent to verbalize the objectives of education based on each cognitive activity level. In 2001, Revised Bloom's Taxonomy (RBT) was introduced by Lorin Anderson and David Krathwohl by revising and developing Benjamin Bloom's concepts into four dimensions: factual, conceptual, procedural, and metacognitive [2]. The number of cognitive domains remains six with changes from noun to the verb form; namely, the domain names have been changed to: Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating [2]. Bloom's Taxonomy has been recommended to be incorporated in both teaching and learning processes, as well as assessment practices [3].

In 2008, the era of Bloom's taxonomy named the Bloom's Digital Taxonomy (BDT) was introduced by Andrew Churches [4]. The BDT consists of six key terms that involve six levels based on difficulty, from the easiest level to the hardest, starting from Remembering and followed by Understanding, Applying, Analyzing, Evaluating, and Creating. Digital taxonomy proposes learning objectives to be in a technology-enhanced teaching context and necessary instructions on how to practice a certain skill to realize its full potential. Each level of taxonomy is supported by an extensive list of thinking skills adapted to a technology-enhanced teaching and learning environment. These activities can make the learning process fun and interesting [5]. The BDT has its own advantages in Bloom's development. The main advantage of the BDT is that educators can get some ideas on how to use digital tools for learning and teaching. Reference [6] emphasized that educators need to be competent in terms of knowledge, skills, and attitudes in learning. One of the main competencies relates to pedagogic competencies. In this work, the BDT is one of the pedagogic components that need to be acquired. However, there is a question of whether teachers use all of BDT activities in their teaching and learning. These findings are crucial for providing information to related parties about current shortcomings in teaching that can be improved to enhance both teaching skills and teachers' confidence in applying digital-based activities in class.

2 Related Work

In recent years, there has been an increasing amount of literature on BDT, which has been published. Some studies have been interested in examining the BDT application to the teaching process. The related studies have shown that the role of teachers in BDT is of great significance when it is used in the classroom. In fact, previous studies have tended to criticize the original Bloom's taxonomy because of its changes toward the digital.

In [7], the knowledge and application of digital verbs and tools by both teachers and students were investigated in order to understand the environment of virtual and conventional learning conceptually. This study has explained that students attending online (distance) learning are better in the application of digital tools, and they can understand them well in addition to their involvement in higher-order thinking tasks, such as publishing and podcasting. Meanwhile, in [1], Benjamin Bloom's psychological and pedagogical model was developed and modified for the systems of adult learning. Based on the analysis, the techniques stemmed from Bloom's taxonomy, which was modified for adult training, enabled the development of students' skills and abilities in analyzing problems thoroughly and comprehensively as well as producing effective and creative solutions.

In [8], the authors described the teaching experience involving a course that introduced educational technologies to teachers in Macau, which was designed based on connectivism, which represented learning theory in the digital era that highlighted the interaction and engagement with digital media and sharing of digital artifacts. The learning outcomes constructively coincided with the activities and assessments of learning and teaching relative to the students' learning needs and disparity in competencies and technological skills, contributing to the discussion regarding how the teachers could learn to teach using digital technologies. In a study conducted in Bulgaria [9], the authors discussed the changes in modern pupils' characteristics, regarding them as a "digital generation," particularly in the area of computer sciences. The results [8,9] showed that introduced changes directly affected the learning-objective taxonomy from classic to revised and ultimately to digital. The dynamics of the learning-objective taxonomy were further explored to clarify the concepts of e-learning, blended learning, electronic learning object, and m-learning based on the digital generation's characteristics. This is vital in order to evolve the learning-objective taxonomy, elucidate the interactivity levels achieved upon the development of e-learning objects to be implemented in the blended learning. It can provide digital instruments as well as authoring tools required by students in the creation of electronic learning objects involving the high cognitive levels of Bloom's digital taxonomy, i.e., evaluation and creation. Generally, the mentioned studies have outlined the teachers' critical role in the implementation of BDT in the teaching process.

In [10], the authors examined the application of Bloom's taxonomy to describe a psychotherapeutic game relative to cognitive processing and knowledge level. The RBT was introduced and applied to five psychotherapeutic games: Personal Investigator, Treasure Hunt, Ricky and the Spider, Moodbot, and SuperBetter. Based on the results, the revised Bloom's Taxonomy was not suitable for comparing the game content. Also, RBT should not be applied to objectively classify psychotherapeutic game content since the results yielded a very low intercoder reliability value.

The adaptation of the revised taxonomy to a new generation of students and a general summary of how Bloom's Original Taxonomy could evolve to Bloom's Revised Taxonomy and initiate Bloom's Digital Taxonomy were presented in [11]. As concluded by the authors, the current restrictions on technology usage in the classroom limited establishment of an association between the classroom and in real life. In [12], it was highlighted that the latest ICT technologies could enhance teaching and learning. The SAMR model and Bloom's taxonomy of educational objectives were implemented at the secondary grammar school for General English and the higher institution for the English for Specific Purposes subject. The two abovementioned studies analyzed the adaptation of Bloom's Original Taxonomy to Bloom's Digital Taxonomy. In [13], thirty compilations posted on websites were analyzed and evaluated on the extent to which these verbs were in alignment with Bloom's taxonomy categories. As explained by the author, Bloom's taxonomy value was heuristic for writing student learning outcomes, and these learning outcomes should be considered by other faculties to describe the expertise level of students who had obtained an associate's, bachelor's, or graduate degree. Reference [14] analyzed the original and revised taxonomies and presented the major criticisms on Bloom's original taxonomy, as well as several criticisms on the revised taxonomy.

Recent taxonomies of objectives and learning-objective strategies can be categorized in terms of the content types (e.g., facts, principles, procedures, concepts, and processes) and performance level (e.g., using and remembering). In [15], a pilot study on the BDT application in an online art project aimed at identifying challenges and affordances in helping amateur artists build their art portfolio through social media sites and other Internet resources was conducted. Due to the high

demand of communication technology and computers in art education, the learning possibilities in the online environment have been required to be extended.

In [16], it has been explained that Bloom's taxonomy can be used by information professionals who train or instruct others how to write learning objectives that describe skills and abilities that learners should master and demonstrate. Since Bloom's taxonomy distinguishes levels of cognitive skill, the Bloom's practice requires learning objectives with a high cognitive skill level and leads to in-depth learning as well as the knowledge and skill transfer to various contexts and tasks. In [17], the responses given by a total of 1,245 science students and 47 science teachers from 14 Catholic high schools in Sydney, Australia, were analyzed. The students and teachers analyzed types of activities using laptops as self-reported, and the BDT was used to differentiate the activities from lower to a higher order. Although the use of pen and paper gradually shifted to using laptops, the students' modal practice entailed the lower-order paradigm of note-taking, as well as working from textbooks electronically using Word processing and electronic textbooks in addition to online searching. In addition, it was observed that students had benefited from higher-order activities, such as blogging and video editing, while teachers were not inclined to engage in these activities.

In Malaysia, only a few studies have discussed Bloom's taxonomy from the digital-based aspect. Previous studies have reported that Bloom's taxonomy has the potential to be applied to different fields, including vocational taxonomic proposals [18], discussion on Bloom improvement in Islamic perspectives [19], promoting creative and critical thinking through English syllabus with augmented taxonomy [20], and understanding more about children's skills in the process of designing digital storytelling games using a tablet [21]. The mentioned studies have highlighted the need for Bloom's taxonomy and BDT to be used in a wider context. However, to the best of the author's knowledge, there have been very limited studies on assessing the psychometric characteristics of BDT measurement items. Quality items can lead to a better measurement of BDT in the local context based on BDT levels. In [18,19], Bloom's Taxonomy was discussed and criticized from the Malaysian context, laying a foundation for further studies on BDT.

In [18], the weaknesses of Bloom's taxonomy in classifying vocational domains were discussed, and a new taxonomy was suggested. Using Delphi techniques, six major domains of vocational taxonomy have identified and verified, namely, knowledge, gross motor skills, fine motor skills, visualization, problem-solving, and inventive skills. Meanwhile, in [19], Western criticism and Islamic views on Bloom's taxonomy were discussed, and it was found that there had been criticisms and improvements from the past studies focusing on Bloom's taxonomy to a new taxonomy regarding four topics, namely, hierarchical arrangements, structural classification, uses, and needs. The new taxonomy refers to the 21st century learning, especially to the field of Islamic measurement context in Malaysia. In contrast, in [20], the relevance of Bloom's taxonomy that included digital elements was discussed from the aspect of augmented reality. A literature review was conducted to examine the extent to which Bloom's Taxonomy of Educational Objectives could be relevant for teaching creative and critical thinking among Malaysian students, identifying the missing aspects in Bloom's taxonomy in the indigenous context as well as highlighting the importance of promoting creative and critical thinking among Malaysian students while reporting the issues surrounding English Literature to be taught as a subject. Finally, the English syllabus, in addition to augmented taxonomy, was suggested based on the outcomes of holistic learning comprising three sets of ability-rationale thinking, purposeful thinking, and context effective relation. However, in this work, one of the objectives is to overcome the critiques by assessing the quality of items for the BDT using the screening list of the existing items without modification, particularly using the new modern measurement theory.

As presented in this section, there have been many studies discussing BDT application to the teaching activities, and some of these studies considered the BDT usage in Malaysia. However, in order to measure the BDT, a very high quality of the measurement items is required. The related literature has emphasized the lack of psychometric assessment of items to measure BDT. This limitation can be overcome using modern assessment theory, such as the Rasch model, to ensure the reliability and validity of the measurement. Hence, the BDT practice application in the classrooms by teachers should be further studied. These findings could provide a useful reference in identifying the teachers' ability to use each level of BDT maximally. In view of all mentioned, one may suggest an approach to examine the psychometric characteristics of BDT measurement items in more detail besides creating a new bloom. Hence, BDT discussion with efforts to provide empirical evidence for new psychometrics items or constructs will help researchers strengthen their future studies more meaningfully.

2.1 Conceptualization and Operationalization

Conceptualization and operationalization used in this study entail BDT based on the definition presented by [4] as shown in Tabs. 2 to 7. The BDT is based on the six key terms, involving six levels based on difficulty, starting from the easiest level that is Remembering, followed by Understanding, Applying, Analyzing, Evaluating, and finally Creating. The BDT definitions are presented in Tab. 1. However, there is a limitation of the use of digital verbs in academia. First, digital tools are grouped based on their appropriate level, which is sometimes difficult to conduct because as they may be used for multiple purposes. In this context, this study examines the extent to which teachers use BDT key terms in the teaching and learning process.

Table 1: Bloom's digital taxonomy definitions

Level	Key term	Definition
Level 1 (L1)	Remembering	Recalling and recognizing knowledge from memory to produce facts, definitions, and lists, as well as reciting or retrieving materials
Level 2 (L2)	Understanding	Developing meaning from different function types verbally or graphically
Level 3 (L3)	Applying	Utilizing or performing a procedure through execution or implementation, involving the use of learned materials through products, such as presentation, models, simulations, and interviews
Level 4 (L4)	Analyzing	Separating concepts or materials into parts to indicate the relation or interrelation between the parts relative to their overall purpose or structure. Mental actions denote the ability to differentiate, organize, attribute, and distinguish between components
Level 5 (L5)	Evaluating	Making decisions by checking and criticizing based on standards and criteria
Level 6 (L6)	Creating	Incorporating the elements for coherent or functional development and regrouping them into a new pattern through planning, generating, or producing

Table 2: Indicators for remembering based on Bloom's digital taxonomy

No.	Remembering	I am doing this in my teaching and learning process to ensure that...
1A	Bullet pointing	Students are able to produce listings in a digital format.
1B	Highlighting	Students are encouraged to select and highlight phrases and keywords as a recalling technique.
1C	Bookmarking or favoriting	Students are able to mark and organize resources, websites, and files for later use.
1D	Social networking	Students are able to establish networks among friends and partners by forging and creating associations between different individuals.
1E	Social bookmarking	Students are able to produce other tags and bookmarks (an online version of local bookmarking or favorites).
1F	Searching or "Googling"	Students are able to simply enter a phrase or keyword into the basic entry pane through search engines.

Table 3: Indicators for understanding based on Bloom's digital taxonomy

No.	Understanding	I am doing this in my teaching and learning process to ensure that...
2A	Advanced and Boolean searching	Students are able to produce, modify, and refine searches according to their needs.
2B	Blog journaling	Students are able to talk, write, or type a daily journal or a task-specific journal to understand the activity report.
2C	Categorizing and tagging	Students are able to organize, structure, and attribute online data and meta-tagging web pages (organizing and classifying websites, files, and materials into folders), as well as understanding the page content to be tagged.
2D	Commenting and annotating	Students are able to make comments and annotations on PDF files, web pages, and other tools, in addition to establishing understandings through comments on pages.
2E	Subscribing	Students are able to subscribe, read, and revisit the subscribe feeds for an in-depth understanding.

2.2 Research Motivation and Aims

In recent times, COVID-19 has been a major public health problem worldwide, including Malaysia, and it has been recording a large number of new cases exceeding a thousand cases daily. The break out of COVID-19 has affected many life aspects, including education. The recent increase in the number of COVID-positive cases has highlighted the need for transforming the teaching and education process from face-to-face methods to online education. The primary concern of this transformation is how to conduct lectures online since this is compulsory for all institutions in order to avoid the risk of further spreading of the COVID-19 virus. The need for online teaching and learning is high since all schools and learning institutions have been closed. However, there are many problems related to online education, such as problems

of unstable Internet access, low student focus, incomplete equipment, and many others. Online learning also makes it difficult for some teachers to assess and test students' achievements and knowledge. Namely, for cognitive assessment, teachers need to be adept at applying BDT since the learning process is conducted online. Thus, teachers should master and use BDT well in their teaching process. To measure the extent to which the BDT aspects are used among teachers, the evaluation of psychometric characteristics on the measurement items is necessary, and it is very important to ensure that the measurements are accurate, especially those involving the use of modern measurement theories.

Table 4: Indicators for applying based on Bloom's digital taxonomy

No.	Applying	I am doing this in my teaching and learning process to ensure that...
3A	Running and operating	Students are able to operate and manipulate hardware and applications in accomplishing a certain objective or goal.
3B	Playing	Students are able to operate or play a game whilst understanding the skill process, task, and application successfully.
3C	Uploading and sharing	Students are able to upload and share materials on sites such as Flickr.
3D	Hacking	Students are able to hack in simpler forms using a simple set of rules to achieve a certain objective or goal.
3E	Editing	Students are able to make editing with the most media (procedure or process employed by the editor).

Hence, this study aims to test whether BDT measurement items have good psychometric characteristics based on the teacher's self-assessment using the Rasch model analysis, which represents modern measurement theory. The modification of assessment by Rasch model will be able to ensure that BDT items difficulty are match with the individual abilities. This study also examines the pattern of the BDT application in teaching from the teachers' perspective.

Table 5: Indicators for analyzing based on Bloom's digital taxonomy

No.	Analyzing	I am doing this in my teaching and learning process to ensure that...
4A	Mashing	Students are able to integrate or mash up a number of data sources into a single resource.
4B	Linking	Students are able to establish/develop associations across and outside web pages and documents.
4C	Reverse-engineering	Students are able to deconstruct to cracking without any associated negative implications.
4D	Cracking	Students are able to crack to comprehend and run the system or application to be cracked as well as analyzing and exploiting its strengths and weaknesses.

Table 6: Indicators for evaluating based on Bloom's digital taxonomy

No.	Evaluating	I am doing this in my teaching and learning process to ensure that...
5A	Blog/Vlog commenting and reflecting	Students are able to comment and evaluate the material in context and reply to postings with constructive criticism and reflective practice.
5B	Posting	Students are able to comment on discussion boards, blogs, and threaded discussions.
5C	Moderating	Students are able to assess comments or postings from various viewpoints in terms of their value, worth, and suitability.
5D	Collaborating and networking	Students are able to effectively collaborate in assessing strengths, abilities, and contributions and have networking, such as engaging and communicating with concerned individuals through a partner network.
5E	Testing (Alpha and Beta)	Students are able to test the processes, applications, and procedures in developing tools by analyzing their purpose or process, correct functions, and their current functions.
5F	Validating	Students are able to affirm the accuracy of their information sources and make judgments by analyzing and evaluating the data sources.

Table 7: Indicators for creating based on Bloom's digital taxonomy

No.	Creating	I am doing this in my teaching and learning process to ensure that...
6A	Programming	Students are able to create programs suitable to their needs and goals (applications, macros, multimedia applications, or games in systematic environments).
6B	Filming, animating, video casting, Podcasting, mixing, and remixing	Students are able to capture, create, mix, and remix content to produce unique products.
6C	Directing and producing	Students are able to view and understand the components to be melded into logical products (production or performance is an extremely highly-creative process in the creation of a product).
6D	Publishing	Students are able to publish not only text but also media or digital formats either from home computers or through the web, which requires an immense overview of the content to be published, as well as the process and products, such as video blogs, blogging, and also wikiing or mashups.

3 Methodology

3.1 Research Design and Sampling

This study adopts a quantitative approach using an online survey research design. The quantitative approach is used because it is suitable for a large number of respondents [22];

this study includes a total of 774 respondents. This approach also helps obtaining more credible findings because it is efficient [22]. Besides, an online survey is very suitable for this study due to the limitations imposed by the COVID-19 pandemic, which has caused difficulty in obtaining research findings in a face-to-face manner. Hence, an online instrument via Google Forms was used. Further, as stated in [23], online surveys have a few advantages such as fast delivery, easy to administer, and inexpensive. Besides, respondents can also answer at their convenience, similar to the mail questionnaire. A stratified sampling technique was conducted on 774 teachers involving five geographical zones in Malaysia, such as North, West, East, South, and Borneo (stratum). Convenient sampling was also used due to the nature of samples that separated the target population into different strata groups. The advantages of such sampling are that it ensures representativeness of samples and estimates the target population with less error and higher precision [22]. Initially, the 200 instruments were given per zone with total of 1000 instruments. However, only 774 were successfully obtained, achieving a return rate of 77.4%. This return rate exceeded the targeted return rate for online collection, which is 60% [24].

3.2 Instrumentation

The measurement originally included 30 measurement items that involved 6 levels of BDT, including 6 items for Remembering (A1 to AF), 5 items for Understanding (2A to 2E), 5 items for Applying (3A to 3E), 4 items for Analyzing (4A to 4D), 6 items for Evaluating (5A to 5F), and 4 items for Creating (6A to 6D). These items were adapted according to the BDT definitions [4]. All of these items were undergone facial validity assessment and content validity by three experts who conducted the Content Validity Ratio (CVR) analysis, as suggested in [25]. The experts were professionals and practitioners, as suggested by [26]. The experts fully agreed to verify item testing and screening, which involved the structure of sentences and words only.

3.3 Data Analysis and Procedure

The data of this study entailed a dichotomous scale that was used to elicit a Yes or No answer [23]. The respondents were required to choose either Agree or Disagree for each item, which indicated the level of agreement, as recommended by [27] as a measurement options. The research included 30 items that were used to identify whether teachers had performed all the key terms for the six BDT levels in their teaching and learning processes. The data were processed using the Rasch model analysis to provide information on: (a) item fit and unidimensionality, (b) Wright map and a bubble chart, (c) mean measure for each BDT level, and (d) reliability and separation index. The WINSTEPS 3.71 was used to perform the Rasch analysis.

The Rasch model assumes that each item comprises only a difficulty parameter and that all items have the same discriminatory index. This ensured that low-capable students could not obtain the correct answer to the items that they did not know by guessing [28]. In short, the probability of success depends on a difference between an individual's ability and the difficulty level of an item. The Rasch model [29] adopts an algorithm that expresses the expectation of the probability of an item as i and the individual's ability as n in the form of a mathematical equation as follows:

$$P\left(X_m = \frac{1}{\theta_n}, \delta_i\right) = \frac{\exp(\theta_n - \delta_i)}{1 + \exp(\theta_n - \delta_i)}. \quad (1)$$

In Eq. (1), $P(X_m = \frac{1}{\theta_n}, \delta_i)$ denotes the probability of respondents n ($n = 1, 2, \dots, N$) with ability β_n to correctly answer an item i ($i = 1, 2, \dots, I$) with difficulty δ_i . This model is regarded as a one-parameter model since the probability P_{ni} denotes the function of difference $(\beta_n - \delta_i)$.

4 Results and Discussion

The results considered several key parameters: (a) item fit and unidimensionality, (b) Wright map and a bubble chart, (c) mean measure by each level of BDT, and (d) reliability and separation index. The obtained results not only showed the quality characteristics of the psychometric items but also indicated the pattern of BDT usage in teaching and supervision from the teachers' perspective. As explained in Section 2.3, there were 30 items assigned with 6 levels of BDT.

4.1 Item Fit

As shown in Tab. 8, 27 out of 30 measurement items fulfilled the fit characteristics in the Rasch model. In Tab. 8, Infit refers to inlier-pattern-sensitive fit statistic and the Outfit refers to outlier-sensitive fit statistic [30].

Table 8: Fit statistics of measurement items

Item	Score total	Measure	Standard error	MNSQ		PTMEA	
				Infit	Outfit	Correlation	Exp.
6A	968	1.61	0.10	1.03	1.18	0.49	0.51
1E	1012	1.20	0.09	1.04	1.13	0.50	0.53
5D	1029	1.05	0.09	0.91	0.89	0.58	0.54
6C	1029	1.05	0.09	0.91	0.89	0.58	0.54
6B	1031	1.04	0.09	0.91	0.95	0.58	0.54
2C	1040	0.96	0.09	1.00	0.95	0.54	0.54
6D	1082	0.62	0.09	1.04	1.03	0.53	0.55
1C	1083	0.61	0.09	0.82	0.79	0.64	0.55
2E	1083	0.61	0.09	0.82	0.79	0.64	0.55
3A	1095	0.52	0.09	0.98	0.98	0.56	0.55
2A	1097	0.50	0.09	1.05	1.06	0.52	0.55
2B	1117	0.35	0.09	0.97	1.08	0.55	0.55
2D	1132	0.23	0.09	0.90	0.86	0.60	0.55
5E	1132	0.23	0.09	0.90	0.86	0.60	0.55
3C	1154	0.07	0.09	0.86	0.82	0.62	0.55
4D	1184	-0.16	0.09	1.06	1.11	0.50	0.54
4C	1196	-0.25	0.09	1.13	1.24	0.46	0.54
4A	1225	-0.46	0.09	1.12	1.17	0.47	0.53
3D	1241	-0.59	0.09	1.07	1.16	0.48	0.53
1D	1288	-0.95	0.09	1.04	1.08	0.48	0.51
3B	1288	-0.95	0.09	1.04	1.08	0.48	0.51
5F	1288	-0.95	0.09	1.04	1.08	0.48	0.51
1A	1304	-1.08	0.09	1.04	1.13	0.47	0.50
4B	1304	-1.08	0.09	1.04	1.13	0.47	0.50
5A	1330	-1.30	0.09	1.01	1.21	0.46	0.49
1B	1345	-1.44	0.09	1.01	1.09	0.47	0.48
5C	1345	-1.44	0.09	1.01	1.09	0.47	0.48
MEAN	1163.8	0.00	0.09	0.99	1.03		
S.D.	114.5	0.90	0.00	0.08	0.13		

Note: Mean square (MNSQ); Point-measure (PTMEA).

The highest measure value was that of item 6A (1.61 logits) and the lowest values corresponded to 5C and IB (-1.44 logits). Overall, activity 6A denoted “programming,” and 5C denoted “moderating.” The standard value of errors was in a range of 0.09–0.10 and complied with the recommended value [31]. Meanwhile, the maximum value of $MNSQ_{infit}$ was 1.13, and the minimum value of $MNSQ_{infit}$ was 0.82. In addition, the maximum value of $MNSQ_{outfit}$ was 1.21, whereas the minimum value of $MNSQ_{outfit}$ was 0.79. Additionally, PTMEA Corr. had a maximum value of 0.64 and a minimum value of 0.46. The range of $MNSQ_{infit}$ was 0.31, and that of the $MNSQ_{outfit}$ range was 0.42, while the range of PTMEA Corr. was 0.18.

Based on the MNSQ fit settings, the used range was from 0.77 to 1.30 [31]. The values that exceeded 1.30 were considered as misfitting, and those less than 0.70 were regarded as overfitting [29]. A total of three measurement items were dropped due to the non-fulfillment of the fit range value, which were: IF ($MNSQ_{outfit} = 1.51$), 5B ($MNSQ_{outfit} = 1.41$), and 3E ($MNSQ_{outfit} = 1.35$). Item IF denoted “searching” or “Googling” by which the students were able to simply enter a phrase or a keyword into the basic entry field of search engines; 5B referred to “posting” by which the students were able to comment on discussion boards, blogs, and threaded discussions; 3E represented “editing” by which the students were able to make editing in most media (procedure or process employed by the editor). The expected score ICC pattern and some unsuitable response patterns (misfits) of items IF, 5B, and 3E are presented by the dotted circle lines in Figs. 1–3, respectively. However, these items were removed because they did not meet the fit requirements.

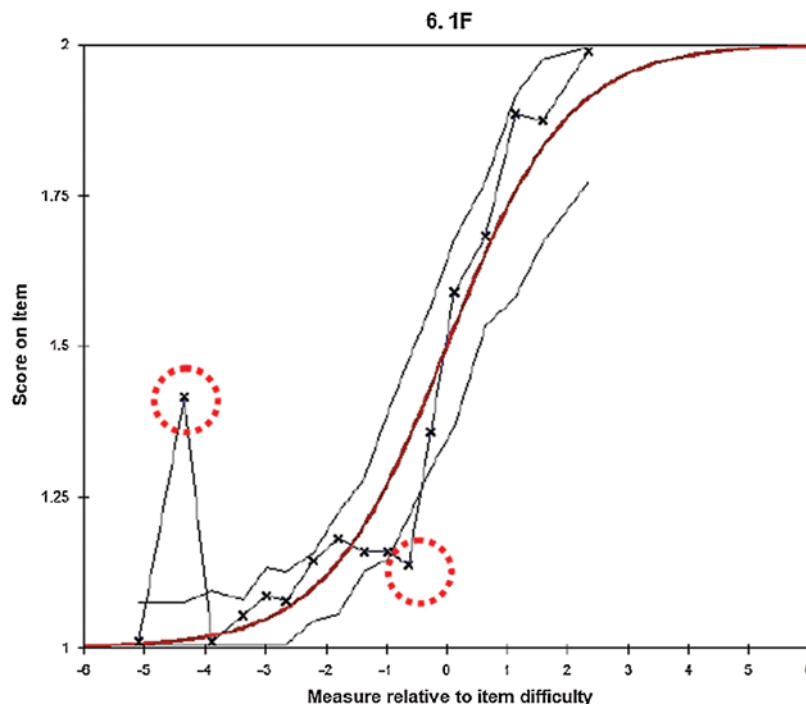


Figure 1: Expected score ICC of 1F

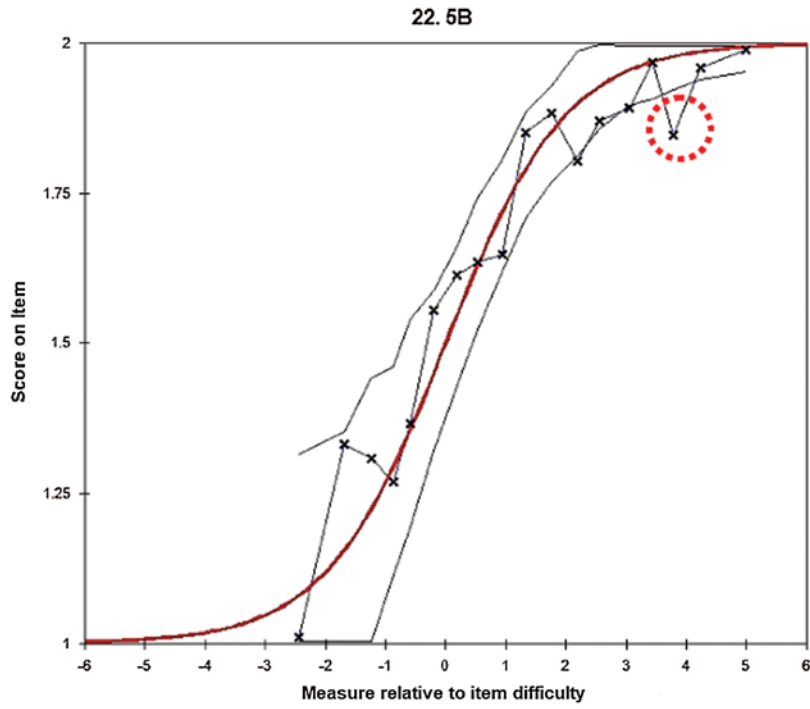


Figure 2: Expected score ICC of 5B

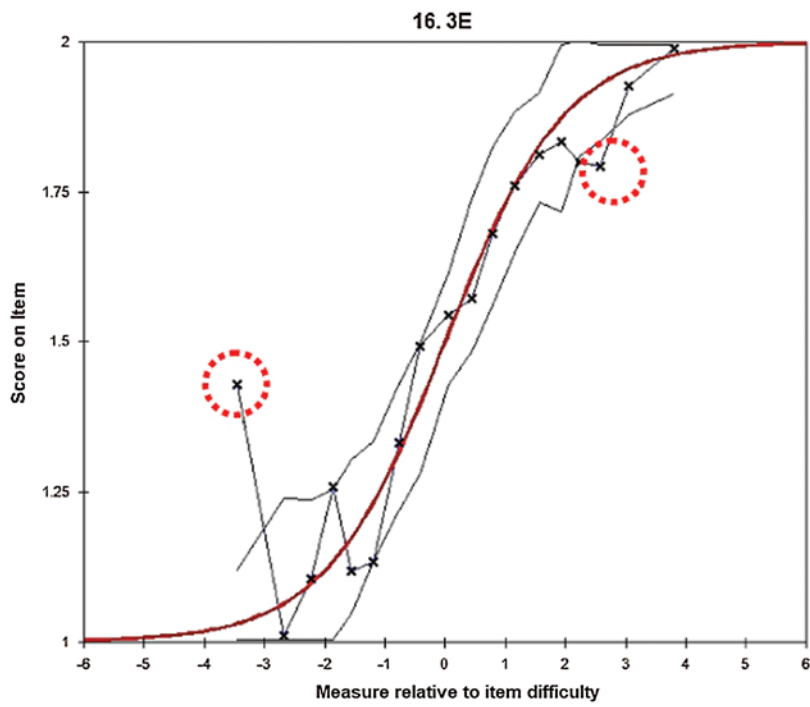


Figure 3: Expected score ICC of 3E

4.2 Items Unidimensionality

The unidimensionality of items indicated that items did not have equality characteristics in the matter to be measured. Dimensionality can be defined as determining an instrument in one direction and one dimension or the force given to one dimension or attribute at once [29] to ensure the instrument's content and construct validity [32]. The raw variance value explained by measures was recorded at 36.3% for overall and each level, which was above the specified value of 20% [33]. The eigenvalue of the entire BDT was 3.5. These values complied with the specified value of less than five [34]. Meanwhile, the overall noise value was recorded at 12.9. These noise values for each L1 to L6 were below 15% [35]. The noise for each construct indicated an underachieved value such as L1 (23.3%), L2 (21.8%), L3 (28.6%), L4 (26.4%), L5 (25.4%), and L6 (29.7%).

4.3 Wright Map and Bubble Chart

Wright Map or item-person map in this analysis denotes a figure that represents items by the item number and the performance of each person to effectively observe the ability of the measured scale items to match the respondents. The distribution of the measurement items according to BDT levels from the aspect of their usage by teachers is presented in Fig. 4.

A total of 55.5% of the total items were above the average difficulty value, while 45% of the total items were below the average difficulty value. This distribution proved that the respondents found it difficult to perform item 6A (Programming), by which the students were able to create programs suitable to their needs and goals (applications, macros, multimedia applications, or games in systematic environments). Meanwhile, the most easily performed activities by the teachers were item IB (Highlighting), by which the students were encouraged to select and highlight phrases and keywords as a recalling technique, and item 5C (Moderating), by which the students were able to assess comments or postings from various viewpoints in terms of their value, worth, and suitability. The results indicated that $item_{max}$ was +1.61, and $item_{min}$ was -1.44. Meanwhile, in the item-person relationship, $person_{max}$ was 3.59, and $person_{min}$ was -4.89. The range values for item and person were 3.05 and 8.48, respectively. The value of μ_{item} was zero, while the value of μ_{person} was 0.05. The mean of individual abilities was slightly higher than the mean of item difficulty, which suggested that BDT measurement items, overall, were easy to perform for the respondents and that, on average, the teachers' performance was higher than the difficulty level of BDT items.

The bubble chart that graphically illustrates the measurement value and item compatibility [36] is presented in Fig. 5. The bubble shape between the overfit and underfit was classified as accepted, which was within the t -value range of ± 2.00 . The bubble chart also shows the bubble positions for all 27 items after the screening was conducted. This screening involved MNSQ because, if MNSQ was considered, Z_{std} could be ignored [34]. The expectation was that difficult items would be answered by more able persons, and easy items would be answered by all. A total of seven items were in the erratic or unpredictable area of two items with Z_{std} value of more than 2.00, which were 4A ($Z_{std\ infit} = 3.0$, $Z_{std\ outfit} = 2.0$) and 4C ($Z_{std\ infit} = 3.2$, $Z_{std\ outfit} = 3.0$). Meanwhile, Z_{std} value of less than 2.00 had five items: 1C ($Z_{std\ infit} = -4.6$, $Z_{std\ outfit} = -2.9$), 2D ($Z_{std\ infit} = -2.6$, $Z_{std\ outfit} = -2.0$), 2E ($Z_{std\ infit} = -4.6$, $Z_{std\ outfit} = -2.9$), 3C ($Z_{std\ infit} = -3.7$, $Z_{std\ outfit} = -2.7$), and 5E ($Z_{std\ infit} = -2.6$, $Z_{std\ outfit} = -2.0$). In this study, erratic or unpredictable referred to items that had Z_{std} value within the Z acceptance range of ± 2.0 , and they were regarded as a misfit.

TABLE 12.2 BLOOM DIGITAL TAXONOMY ZOU482WS.TXT Nov 26 22:54 2020
 INPUT: 774 TEACHER 30 item REPORTED: 774 TEACHER 27 item 2 CATS WINSTEPS 3.71.0.1

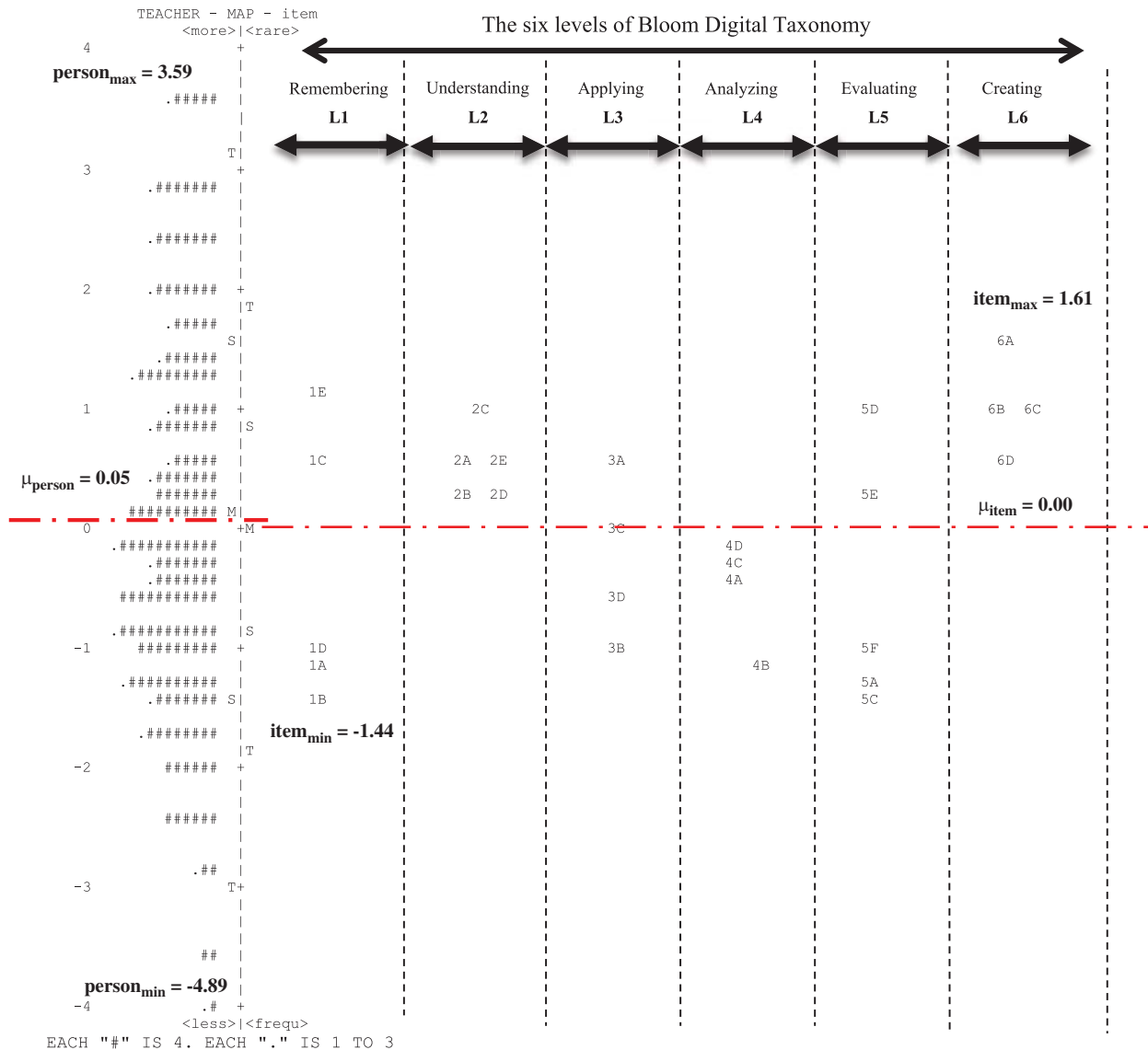


Figure 4: Wright item-person map

The L1 analysis (Remembering) showed that item IE was the hardest item with 1.20 logits. This result showed that Social Bookmarking was the least applied activity by teachers in the teaching and learning process at the L1 level. Meanwhile, activity 1B (Highlighting) included the most performed items by teachers with -1.44 logits. The L2 analysis (Understanding) showed that item 2C was the hardest item with 0.96 logits. This result indicated that Categorizing and Tagging was the least applied activity by teachers in the teaching and learning process. In contrast, activity 2D (Commenting and Annotating) included the most performed items by teachers with 0.23 logits. The L3 analysis (Applying) showed that item 3A was the hardest item with 0.52 logits,

indicating that Running and Operating was the least applied activity by teachers in the teaching and learning process. Meanwhile, activity 3B (Playing) included the most performed items by teachers with -0.95 logits. The L4 analysis (Analyzing) showed that item 4D was the hardest item with -0.16 logits; thus, Cracking was the least applied activity by teachers in the teaching and learning process at the L1 level. However, activity 4B (Linking) included the most performed items by teachers with -1.08 logits. The L5 analysis (Evaluating) showed that item 5D was the hardest item with 1.05 logits; Collaborating and Networking was the least applied activity by teachers in the teaching and learning process. Meanwhile, activity 5C (Moderating) included the most performed items by teachers with -1.44 logits. The L6 analysis (Creating) showed that item 6A was the hardest item with 1.61 logits, and Programming was the least applied activity by teachers in the teaching and learning process. Last, activity 6D (Publishing) included the most performed items by teachers with -1.44 logits 0.62 logits.

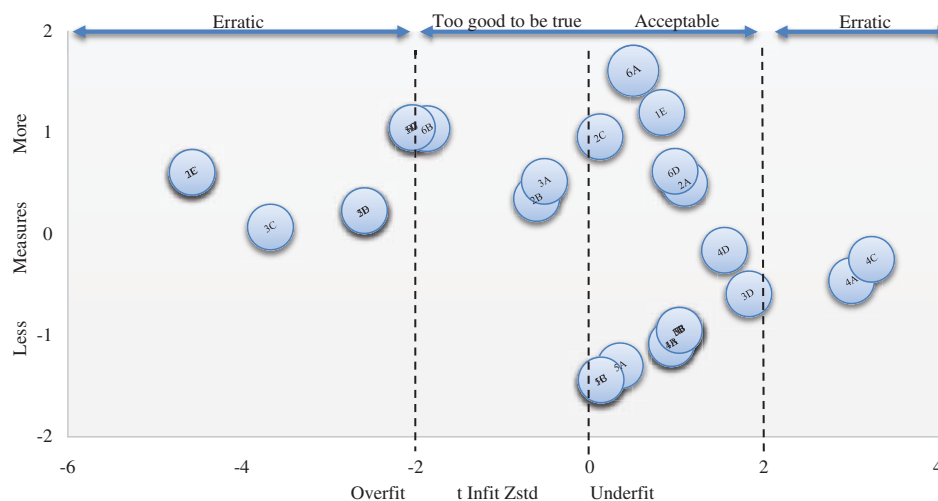


Figure 5: Bubble chart of item fitness

4.4 Mean Measure of Each BDT Level

Based on the mean logit value results, the hardest level of BDT to be conducted by teachers in their teaching and learning was Level 6, which was Creating ($+1.08$), followed by Level 2 ($+0.53$), Level 3 (-0.24), Level 1 (-0.33), and Level 5 (-0.48), while the easiest level was Level 4, i.e., Analyzing. Based on the level of BDT, Level 6 (Creating) is the highest level of BDT. This level represents the hardest level from the aspect of activity implementation. Thus, the results presented in this study are logical because the logit value of 1.08 has indicated level L6 as the most difficult activity for teachers to implement. However, interestingly, the results showed that the easiest level to be implemented by the teachers was not L1 (remembering) as it was expected but L4 (analyzing). According to [4], L4 can be defined as separating concepts or materials into parts to determine the relation or interrelation between the parts relative to their overall purpose or structure. Level L4 also includes mental actions, which comprise the ability to differentiate, organize, attribute, and distinguish between components.

This result could be caused by the elements of Higher Order Thinking Skill (HOTS) instilled by teachers to the students. The Malaysian Education Development Plan 2013–2025 explains

that national examinations and school-based assessments (PBS) have been revamped to gradually increase the percentage of questions that define high-level thinking skills. By 2016, high-level thinking questions included at least 40% of the questions in *Ujian Penilaian Sekolah Rendah* (UPSR), and at least 50% of the questions in *Sijil Pelajaran Malaysia* (SPM). More group-based projects and assignments were also done to improve students' high-level thinking skills and their ability to work individually and in groups. They were given more community-based projects and cross-school activities to foster interaction between individuals from all backgrounds. In addition, Wave 1 (2013–2015) played an important role in changing the education system by supporting teachers focusing on key skills and redesigning exam questions to put a higher focus on high-level thinking skills questions [37]. The high-level thinking should be starting at Analyzing level; therefore, the findings indirectly had proven that the implementation of the BDT activities by teachers were at a higher level.

4.5 Reliability and Separation Index

The rating scale instrument quality criteria used in this study are based on the setting in [31]. Person reliability entails the consistency of person ordering to be accepted under conditions that the equivalent set of items that measures the same construct is given to this respondents [38]. The overall person measurement reliability value was 0.87. For all levels of persons ability, which could be considered as good. While the item measurement reliability was 0.99, which could be regarded as excellent; the person separation index was 2.60, which could be considered as satisfactory. Item separation denotes the ability of all participants to answer all the items' difficulty levels. This means that the respondents can be distinguished by the constructs being tested [38]. The item separation index was 9.74, which could be considered as excellent. This means that the quality of BDT measurement items in this instrument is excellent, but the consistency of answers from teachers is only fair. The grouping of persons and items can be obtained using the following formula: $H = [(4 * Separation) + 1]/3$, where H represents the separation value, which can be taken from reliability and separation index produced by the Winsteps software.

The overall separation value of persons was $H_{person} = (4 * 2.6 + 1)/3 = 3.8$, which could be rounded to 4. This means that there were four groups of teachers according to the ability levels. These findings hold for all BDT levels, from L1 to L6. Meanwhile, the overall separation value of items was $H_{item} = (4 * 9.74 + 1)/3 = 13.32$, which could be rounded to 13. This means that there were 13 clusters of items according to the difficulty levels. The H values of different levels were as follows: $H_{item(L1)} = (4 * 11.15 + 1)/3 = 15.2$; $H_{item(L2)} = (4 * 2.63 + 1)/3 = 3.84$; $H_{item(L3)} = (4 * 6.30 + 1)/3 = 8.73$; $H_{item(L4)} = (4 * 3.83 + 1)/3 = 5.44$; $H_{item(L5)} = (4 * 10.45 + 1)/3 = 14.27$; and $H_{item(L6)} = (4 * 3.60 + 1)/3 = 5.13$. These results indicated that the items of level L1 could be grouped into 15 difficulty levels, items of level L2 could be grouped into four difficulty levels, items of level L3 could be grouped into nine difficulty levels; items of level L4 could be grouped into five difficulty levels; items of level L5 could be grouped into 14 difficulty levels, and finally, items of level L6 could be grouped into five difficulty levels. In comparison to [29], the separation values with more than two levels are sufficient. In summary, these results help to improve the item quality provided to teachers during self-assessment. The findings of this study will directly benefit teachers with lacking of BDT in regaining the digital element for teaching. The Rasch psychometrics evidence may be help the researcher to measure BDT accurately. This will enable teachers that have struggle to implement the BDT to be more dynamics and creative in teaching pedagogically.

5 Conclusion and Future Works

The current study aims to improve the measurement for BDT items through teachers' self-assessment in teaching and learning, and the Rasch measurement model is proposed for the assessment of psychometric properties. The results show that a total of 27 measurement items can be used as an alternative for BDT measurement using the Rasch model. The results show that the Rasch model can more clearly demonstrate various item properties compared to the classical test theory. Moreover, this study indirectly shows to which extent teachers tend to apply each level of BDT in their teaching and learning practice and examines which BDT activities are the hardest and easiest to apply. However, certain limitations need to be considered in future works. First, the results presented in this study are applicable only to the Malaysian population, so the study should be expanded regarding both contexts and countries. Namely, it would be interesting to explore and compare more characteristics of item response for various levels of respondents' ability through systematic comparisons. Second, this investigation has been limited to the teachers' perspective, so future research is highly encouraged to introduce scale analysis to develop specific questionnaires from the perspective of students' understanding of the BDT levels. In fact, this measurement construct can be tested for its validity using multivariate analysis, such as factor analysis or principal component analysis, to provide empirical evidence for future reference. Third, this study is limited to general definitions in each level, starting from L1 to L6, so further investigations can be performed to each specific activity for every level of BDT. The information can be useful for customizing digital teaching activities that suit both teacher and students' abilities relative to the implementation of teaching and learning in class.

Acknowledgement: The author would like to express the utmost gratitude to the anonymous reviewers for their insightful comments and suggestions, which helped to improve the content, quality, and presentation of this research paper. The author would also like to honor and give high appreciation to the Faculty of Education, University Kebangsaan Malaysia (UKM).

Funding Statement: This study was funded by the Ministry of Higher Education (MOHE), Malaysia under the Fundamental Research Grant Scheme (FRGS) (FRGS/1/2018/SSI09/UKM/02/1), and the Faculty of Education, University Kebangsaan Malaysia (UKM), Geran Galakan Penyelidik Muda (GGPM-2017-088).

Conflicts of Interest: The author declares that he has no conflicts of interest to report regarding the present study.

References

- [1] N. Grebin, S. Grabovska, R. Karkovska and A. Vovk, "Applying Benjamin Bloom's taxonomy ideas in adult learning," *Journal of Education Culture and Society*, vol. 11, no. 1, pp. 61–72, 2020.
- [2] D. R. Krathwohl, "A revision of Bloom's Taxonomy: An overview," *Theory into Practice*, vol. 41, no. 4, pp. 212–218, 2002.
- [3] M. T. Chandio, S. M. Pandhiani and S. Iqbal, "Bloom's taxonomy: Improving assessment and teaching-learning process," *Journal of Education and Development*, vol. 3, no. 2, pp. 203–211, 2016.
- [4] A. Churches, "Bloom's digital taxonomy," pp. 1–44, 2008. [Online]. Available: <http://burtonslifelearning.pbworks.com/f/BloomDigitalTaxonomy2001.pdf>.
- [5] A. Gani, "Improving the ability of implementing fun-based learning through structured clinical supervision on Mathematics teachers at SMA Negeri 1 Barru Indonesia," *Journal of Applied Science, Engineering, Technology, and Education*, vol. 3, no. 2, pp. 195–202, 2020.

- [6] J. J. Pongkendek, D. N. Marpaung, D. S. Ahmar and S. Rahmatia, "The professional competence analysis of chemistry teacher of senior high school in Merauke," *Journal of Applied Science, Engineering, Technology, and Education*, vol. 3, no. 1, pp. 46–52, 2020.
- [7] H. Amin and M. S. Mirza, "Comparative study of knowledge and use of Bloom's digital taxonomy by teachers and students in virtual and conventional universities," *Asian Association of Open Universities Journal*, vol. 15, no. 2, pp. 223–238, 2020.
- [8] N. Thota and J. G. M. Negreiros, "Introducing educational technologies to teachers: Experience report introducing educational technologies to teachers: Experience report," *Journal of University Teaching & Learning Practice*, vol. 12, no. 1, pp. 1–13, 2019.
- [9] E. Goranova, "Creation of electronic learning objects for the high cognitive levels of Bloom's digital taxonomy," *Knowledge International Journal*, vol. 31, no. 2, pp. 585–590, 2019.
- [10] P. Haring, H. Warmelink, M. Valente and C. Roth, "Using the revised Bloom taxonomy to analyze psychotherapeutic games," *International Journal of Computer Games Technology*, vol. 1, no. 1, pp. 1–9, 2018.
- [11] B. C. Wedlock and R. Grove, "The technology driven student: How to apply Bloom's revised taxonomy to the digital generations," *Journal of Education & Social Policy*, vol. 4, no. 1, pp. 25–34, 2017.
- [12] J. Netolicka and I. Simonova, "SAMR model and bloom's digital taxonomy applied in blended learning/teaching of general english and ESP," in *Int. Symp. on Educational Technology 2017, ISET 2017, Proc. IEEE Computer Society Conf. Publishing Services*, Hong Kong, China, pp. 277–281, 2017.
- [13] C. J. Stanny, "Reevaluating bloom's taxonomy: What measurable verbs can and cannot say about student learning," *Education Sciences*, vol. 6, no. 4, pp. 1–12, 2016.
- [14] S. M. A. Soozandehfar and M. R. Adeli, "A critical appraisal of Bloom's taxonomy," *American Research Journal of English and Literature*, vol. 2, no. 1, pp. 1–9, 2016.
- [15] S. Lee, "Pilot study: Bloom's digital taxonomy application for an online art project," *Literacy Information and Computer Education Journal*, vol. 7, no. 2, pp. 2268–2274, 2016.
- [16] N. E. Adams, "Bloom's taxonomy of cognitive learning objectives," *Journal of the Medical Library Association*, vol. 103, no. 3, pp. 152–153, 2015.
- [17] S. J. Crook and M. D. Sharma, "Bloom-ing heck! the activities of Australian science teachers and students two years into a 1:1 laptop program across 14 high schools," *International Journal of Innovation in Science and Mathematics Education*, vol. 21, no. 1, pp. 54–69, 2013.
- [18] N. L. Long and R. Mustapha, "Bloom's taxonomy analysis in vocational assessment: The development of a new Taxonomy using the Delphi Technique," *Journal of Quality Measurement and Analysis*, vol. 15, no. 1, pp. 65–75, 2019.
- [19] N. W. Masrom, M. Hashim, N. Hashim and F. P. Behak, "Bloom's taxonomy position from an Islamic perspective," *Journal of Quran Sunnah Education and Special Needs*, vol. 2, no. 1, pp. 18–26, 2018.
- [20] S. A. Rahman and N. F. A. Manaf, "A critical analysis of Bloom's Taxonomy in teaching creative and critical thinking skills in Malaysia through english literature," *English Language Teaching*, vol. 10, no. 9, pp. 245–256, 2017.
- [21] M. Hayati and M. Yatim, "Exploring 21st century children's skills in the process of digital storytelling game design," *EDUCATUM Journal of Science, Mathematics and Technology*, vol. 2, no. 1, pp. 82–96, 2015.
- [22] J. F. Hair, M. W. Celsi and D. E. Harrison, *Essentials of Marketing Research*, 5th ed. New York, USA: McGraw-Hill Education, 2020.
- [23] U. Sekaran and R. Bougie, *Research Methods for Business: A Skill Building Approach*, 7th ed. New Delhi, India: John Wiley & Sons, 2016.
- [24] J. E. Fincham, "Response rates and responsiveness for surveys, standards, and the journal," *American Journal of Pharmaceutical Education*, vol. 72, no. 2, pp. 2–4, 2008.
- [25] C. H. Lawshe, "A quantitative approach to content validity," *Personnel Psychology*, vol. 28, no. 4, pp. 563–575, 1975.

- [26] D. M. Rubio, M. Berg-Weger, S. S. Tebb, E. S. Lee and S. Rauch, "Objectifying content validity: Conducting a content validity study in social work research," *Social Work Research*, vol. 27, no. 2, pp. 94–105, 2003.
- [27] A. DeCastellarnau, "A classification of response scale characteristics that affect data quality: A literature review," *Quality and Quantity*, vol. 52, no. 4, pp. 1523–1559, 2018.
- [28] A. A. Aziz, *Rasch Model Fundamentals: Scale Construct and Measurement Structure*, Kuala Lumpur, Malaysia: Integrated Advance Planning Sdn Bhd, 2011.
- [29] T. G. Bond and C. M. Fox, *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, Third Avenue, NY, USA: Routledge, 2015.
- [30] J. M. Linacre, "What do infit and outfit, mean-square and standardized mean," *Rasch Measurement Transactions*, vol. 16, no. 2, pp. 878, 2002.
- [31] J. W. P. Fisher, "Rating scale instrument quality criteria," *Rasch Measurement Transactions*, vol. 21, no. 1, pp. 1095, 2007.
- [32] M. Wu and R. Adams, *Applying the Rasch Model to Psycho-Social Measurement: A Practical Approach*, Melbourne, Australia: Educational Measurement Solutions, 2007.
- [33] D. Reckase, "Unifactor latent trait models applied to multifactor tests: results and implications," *Journal of Educational Statistics*, vol. 4, no. 3, pp. 207–230, 1979.
- [34] J. M. Linacre, *A User's Guide to WINSTEPS: Rasch Model Computer Programs*, Illinois, USA: MESA Press, 2005.
- [35] A. A. Aziz, M. S. Masodi and A. Zaharim, *Rasch Measurement Model Basis: Scale Development and Measurement Structure*, 3rd ed. Bangi, Malaysia: UKM Press, 2017.
- [36] J. M. Linacre, *A User's Guide to WINSTEPS: Rasch Model Computer Programs*, Illinois, USA: MESA Press, 2012.
- [37] Malaysia Ministry of Education, "Malaysia Education Development Plan 2013–2025 (Preschool to Post-Secondary Education)," Putrajaya, Malaysia: Ministry of Education, 2013. [Online]. Available: <https://www.moe.gov.my/muat-turun/penerbitan-dan-jurnal/1818-pelan-pembangunan-pendidikan-2013-2025/file>.
- [38] B. D. Wright and G. N. Masters, *Rating Scale Analysis Rasch Measurement*, Illinois, USA: MESA Press, 1982.