PRACTICALITY TEST OF "STEPPING THE 5 STAIRS A TECHNIQUE" ON GRAMMAR FOR COLLEGE STUDENTS: RASCH MODEL ANALYSIS

Satria Agust[*], Gatot Subroto

Universitas Maritim Raja Ali Haji, Kota Tanjungpinang, Kepulauan Riau 29115, Indonesia

Abstrak

Menyingkap data penelitian dengan menggunakan aplikasi yang tidak dapat dipercaya membawa temuan-temuan penelitian diambang kehancuran. Tidak seperti aplikasi-aplikasi yang alinnya, aplikasi Rasch model bisa menganalisa data penelitian apa adanya. Aplikasi ini juga bisa menghilangkan jegaduhan dan keraguan para peneliti. Aplikasi ini digunakan untuk menganalisa instrumen penelitian pada penelitian pengembangan teknik pengajaran grammar untuk mahasiswa. Tujuan penelitian ini adalah untuk menganalisa instrument uji kepraktisan yang digunakan dalam penelitian pengembangan teknik grammar, "Stepping the 5 Stairs A" untuk mahasiswa dengan menggunakan Rasch model. Penelitian ini adalah kuantitatif deskriptif. Tempat penelitian di sebuah universitas di Provinsi Kepulauan Riau dan dilakukan pada bulan Mei 2018. Mahasiswa tahun pertama jurusan Pendidikan Bahasa Inggris Tahun Akademik 2017/2018 menjadi subyek penelitian. Penelitian ini menghasilkan bahwa analisis instrumen uji kepraktisan dengan menggunakan Rasch model menunjukan rata-rata di atas item logit (+0.00 logit) yang menekankan persetujuan partisipan pada kualitas intrumen tersebut. Cara pandang yang berbeda yang partisipan berikan melalui pengisian instrumen menjelaskan keadaan mereka saat ini yang perlu untuk diteliti lebih lanjut di dalam penelitian ini.

Kata Kunci: Rasch model, stepping the 5 stairs A, instrument uji kepraktisan

Abstract

Exposing research data by using unreliable applications bring about research findings to the brink of collapse. Unlike other applications, Rasch model application can analyze research data as it is. It can also dismiss researchers' hubbub and anxiety. This application is used to analyze research instrument on development of a teaching technique on grammar for college students. The objective of this research is to analyze practicality test instrument used in research and development on grammar technique, "Stepping the 5 Stairs A" for college students by using Rasch model. This research was descriptive quantitative. The setting of this research was in a university in Riau Isle and was conducted in May 2018. The first-year students of English Education in 2017/2018 academic year became the research subject. This research results that the analysis of the practicality test instrument by using Rasch model shows the above average logit item (+0.00 logit) emphasizing participants' approval on the quality of the instrument. Different perspectives which participants give through the instrument fulfilment describe their present conditions which are necessary to have further investigation in this research.

Key words: Rasch model, stepping the 5 stairs A, practicality test

1. Introduction

Working on social researches in education demands guaranteed applications. The applications can help researchers find it easy to analyze research data. They can also bring the researchers' works to high reputable publications on national or international level journals.

Unfortunately, not all of them can give accurate analyses without manipulations of research data which are able to result in shortcomings on the researchers' works. Rasch model can analyze research data as it is without manipulation because it serves equal probability on data being analyzed. Besides, it can be used to analyze

*Penulis Korepondensi
Email Address : satria.agust@umrah.ac.id

either dichotomous or polytomous research instruments. The instruments analyzed using this application will be valid, reliable, accurate, and guaranteed, whereas using other applications is still questionable. The anxiety of the researchers using their research data analyses of Rasch model will vanish.

Some comparative works which involved CTT and Rasch model have been conducted. The results are different. Some say that they find stunning similarities of their measurements. Some others say that the measurements of CTT and IRT on person and item statistics were very comparable. It seems that there is dubiousness on the experts' perspective of the two models. It is not about the measurement result but it tends to a great deal of change that is made by the two models. (Stage, 2018) compared those two models on the Swedish Scholastic Aptitude Test (SweSAT). She found that the IRT did not fit the test data, while the CTT could do better than that of the IRT. She continued that a positive assumption of IRT models is unidimensionality, meaning that the measurement of items in a test is one single ability. Some models could measure more than one ability called as multidimensional models and they are more complex and are still not well-developed as the unidimensional models.

The CTT views item difficulty and item discrimination as sample dependent. It classifies higher item difficulty through test-taker samples of lower-average knowledge and lower item difficulty through test-taker samples of above-average knowledge. In case of discrimination, higher scores can be obtained through heterogeneous test-taker samples and lower ones could be found in homogeneous samples. The relationships of sample dependency cut down the quality of the overall utility of these statistics (Schumacker, 2018). The weak assumptions and conceptual simplicity of the CTT have affected the lack of refinement in many legacy measures. For example, the CTT model uses ordinal scale, whereas IRT scaling deals with interval one (Kean & Reilly, 2018). To some degree, the IRT models offer another advantage i.e. representing

the test-takers' ability and the items of difficulty as independent parameters, whereas the CTT dealing with sample-specific in its values cannot separate these two constructs. Unlike the CTT, the IRT has a way to separate these two empirically plaited concepts which is known that no other psychometric model can do (Mead & Meade, 2018). There are some constraints which the TCC has. They are (1) item difficulty and item discrimination depending on particular examinee samples; (2) reliability definition in CTT established through the concept of parallel tests; (3) an assumption of standard error of measurement which is the same for all subjects; and (4) focusing on level information of test toward the exclusion of item level information (OoN, 2018).

The researchers will not join the gaps between different perspective of the two models. They just want to focus on the instrument used in their development research. The instrument is practicality test. The practicality test instrument is an on-line distributed instrument by using google form. The instrument was used to know the students' responses of how their points of view after being taught by using "Stepping the 5 Stairs A" technique on grammar. This instrument was developed by the researchers themselves. The instrument was distributed through the link which the students could come to the access to fulfil each item. The reliability test of the instrument was obtained through the Rasch model analysis. Besides, Stepping the 5 Stairs A technique is a teaching technique on grammar for college students which has been being developed. This is a new technique and not a modified teaching technique which can help students construct English sentences accurately, correctly, and efficiently. A letter "A" means that the technique focuses on full verb sentences.

This research was aimed at knowing the Rasch model analysis on the practicality test instrument used in the development research of a teaching technique, "Stepping the 5 Stairs A", on Grammar for College Students. The comparative studies between IRT and classical test theory

(CTT) are coming up to surface. The two models of instrument analyses have become current debate to argue "we are the best". The researchers do not want to compare the two models. They just want to implement the Rasch model to analyze the practicality test instrument which uses the interval scale instead the ordinal one and which can analyze every single item in the instrument. It can give them valuable information to conduct further research in the development of a teaching technique, "Stepping the 5 Stairs A".

2. Method

This research used a descriptive quantitative approach (Creswell, 2014) and conducted at a university in Riau Archipelago and it was done in May 2018. The participants involved in this research were the first-year students of English Education Department, consisting of twenty-two students. However, there were only twenty students who were involved in during the research for the other two students had problem to fulfil the requirement of the minimum presence in a class. The researchers used the purposive sampling technique on this research.

The researchers analyzed the practicality test questionnaire used in the development research "Stepping the 5 Stairs A technique". The questionnaire was fulfilled through on-line system by using google form to collect data. The practicality test data were analyzed by using Rasch model software called WINSTEPS version 3.73. The Cronbach-alpha (KR20), in the summary statistics, shows us on the questionnaire reliability. The Cronbach-alpha value for the questionnaire was .90 which meant that the instrument is excellent and acceptable (Sumintono & Widhiarso, 2013).

3. Result and Discussion

From the summary statistics of the questionnaire, it shows that the score of the measured person in the questionnaire used in development research of teaching technique, "Stepping the 5 Stairs A", on Grammar for College Students was +2.24 logit. It is higher than the requirement of the minimum logit score i.e. 0.0. It means that there was a tendency of the participants to choose "agree" with almost of all items of the questionnaire. The measured person on INFIT MNSQ score was .95. Furthermore, the score of OUTFIT MNSQ was 1.09. Because the above obtained scores almost reached the ideal score i.e. 1.00, it means that the quality of the questionnaire was good. For the INFIT ZSTD, the score was -.2, whereas the score of OUTFIT ZSTD was -.1. When the ideal score of the INFIT ZSTD and OUTFIT ZSTD i.e. 0.0 was compared to the obtained scores, it means that the quality of the questionnaire was also good. Besides, the score of the person reliability was .82, meaning that the instrument used in this development research was good in category. The person separation score was 2.11. The other formula the so-called "H" was used to see the grouping thoroughly and it is called the separation level. The obtained score of it was 3.11 and it was rounded below i.e. 3, which means that there were three groups of participants within this development research.

The Cronbach Alpha (kr-20) score was .90. It was used to measure the interaction between person and item. The obtained score indicates an excellent interaction. The score of the reliability test of the measured item was .83. It means that the questionnaire was good in the category and the instrument used in the development research corresponded to the requirement of the Rasch model. In addition, the measured item score of the INFIT MNSQ of the item table was .96 and the score of the OUTFIT MNSQ was 1.09. For the INFIT ZSTD of the measured item, the score was -.1, whereas the score of OUTFIT ZSTD was .1. Through the obtained scores of MNSQs and ZSTDs, it means that the quality of the questionnaire was also good. The item separation score of the measured item was 1.23. The obtained score by using the "H" formula was 1.97 and it was rounded above to be 2 which means that the questionnaire having been used in the development research showed the students' ability separation.

The research subject was the college students in a university located in Riau Archipelago province comprising 20 students in total. They

were from different backgrounds and from different islands existing in Riau Isle. The researchers were bit surprised after getting information about the students' reasons joining the English Education Study Program i.e. that was not their passion. There was 55 percent giving reason not interested in taking the major. The other 40 percent participants claimed that they put their interests on English Education major but there was 5 percent showing no passion on English Education but trying to survive, and showing a good progress in processes of learning. Table 1 highlights information details relating to the participants' backgrounds.

Table 1. Participants' Demographic Information

Demographic information	Total number	Percentage
Sex		
Male	4	20%
Female	16	80%
Total	**20**	**100%**
Origin		
Rural	10	50%
Urban	10	50%
Total	**20**	**100%**

Instruments measurement; through the Wright map analysis result with the person map of the questionnaire in figure 1, it shows that the 20 participants had different perspectives in answering the questionnaire items. The right side is for the person column and the left one is for the item column. The items which were above the item logit average (+0.00 logit) means that the participants found it difficult to agree while those which were below the item logit were considered as easy to be accepted. Figure 1 also describes that the third item was the most difficult item to agree almost all of the participants, whereas the fifth and the sixth items were the easiest items to agree almost all of the participants.
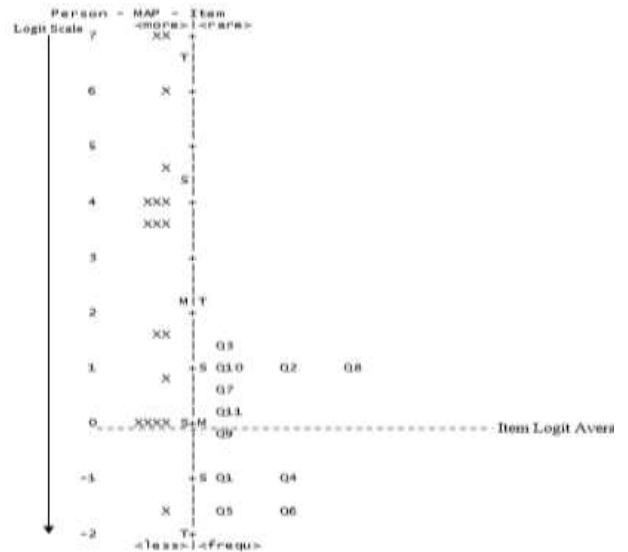

Figure 1. Wright Map of questionnaire items

Figure 2 explains that as mentioned before the obtained score of the "H" formula of the questionnaire the measured person was 3, it means that there were three different levels of participants seen from their origin i.e. 'strongly agree', 'agree', and 'disagree'. The participants who are from rural areas surrounding Riau Archipelago chose 'disagree' in the questionnaire for 38.89%, while those who are from urban areas got 16.67%. Furthermore, the participants who are from rural areas surrounding Riau Archipelago chose 'agree' in the questionnaire for 11.11%, while those who are from urban areas got 27.78%. The participants who are from rural areas surrounding Riau Archipelago did not choose 'strongly agree' in the questionnaire while there was one participant from urban areas chose 'strongly agree' for 5.56%. These three different groups correspond to the obtained "H" formula score through the person separation taken form the summary statistics of Rasch model.

Figure 3 also describes that the rural participants found it difficult to choose 'strongly agree' or 'agree' with the items 1, 7, 8, and 10 and found it easy to come to 'strongly agree' or 'agree' with the items 1, 5, and 6. Besides, this figure also deciphers the different items which the urban participants chose. The items 2, 3, 8, and 10 were hard to be chosen as 'strongly

agree' or 'agree' for them. They even found it easy to choose 'strongly agree' or 'agree' with the items 3, 4, and 5. It can be concluded that those coming from rural and urban areas surrounding Riau Isle had the different perspectives on responding the questionnaire. Their localities affect them to correspond to certain items of the questionnaire used in development research of teaching technique, "Stepping the 5 Stairs A", on Grammar for College Students.
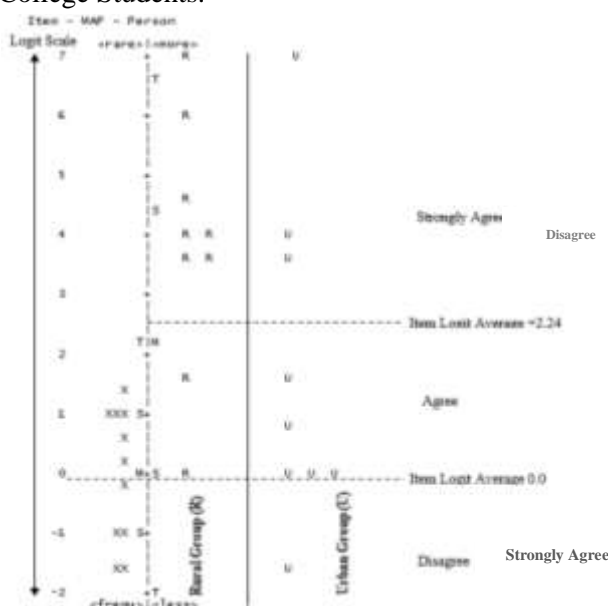

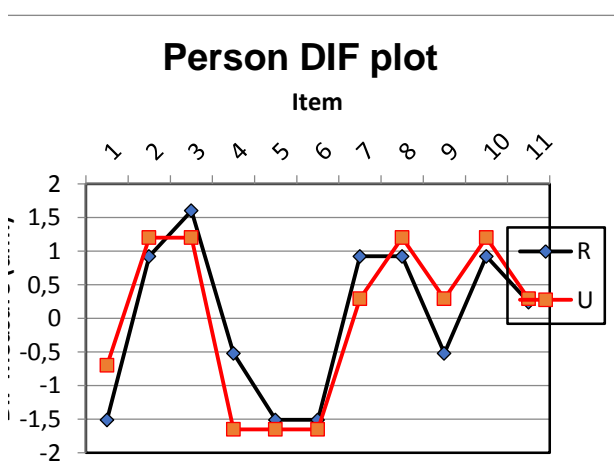**Figure 2.** Wright Map based upon the participants' origin toward the questionnaire

Figure 3. Person Different Plot on questionnaire

As mentioned previously, applying Rasch model to analyze the practicality test instrument made the researchers easy to measure the reliability of the instrument. It is parallel with (Chan, Ismail, & Sumintono, 2014) who state that Rasch measurement model can give us information relating to the person and item reliability. Through the statistical summary as previously mentioned, it was known that the value of Cronbach-alpha for the reliability of the practicality test instrument used in this research was .90, person reliability value was .83, and item reliability value was .60. It can be said that the reliability of the questionnaire was excellent in category, while students' consistency in answering the questionnaire was good in category, and the quality of the questionnaire items was enough in category (Sumintono & Widhiarso, 2015). In this way, the researchers can say that Rasch model describes a precise analysis of the instrument used in the development research "Stepping the 5 Stairs A technique. The ability to identify the misfitness of the items and respondents of the research is one of the merits of Rasch modeling method (Ariffin, Omar, Isa, & Sharif, 2010). Moreover, it is said that Rasch model is easy and effective to be used in analyzing, explaining, and summarizing the data of the research gathered from questionnaire or test (Sumintono & Widhiarso, 2013).

4. Conclusion

To sum up of the Rasch model role in analyzing the practicality test instrument used in this research is incredible. It can analyze the instrument used in this research as needed and precisely. The statistical summary of the instrument reveals that the quality of the instrument was good in category. Having gained the valid and reliable research instrument, the researchers can come to the next steps of the research on the development of a teaching technique, "Stepping the 5 Stairs A", on grammar for college students. Besides, another merit which Rasch model can serve through this research is that it can analyze research instrument based upon the researchers' needs. Through the use of the Rasch model, the researchers not only obtain the reliable research instrument but also grasp value information of

how the participants involved in this research give response and correspond to the on-line distributed instrument. The demography information like sex and origins can also be analyzed. The analysis on sex and origins is to know whether or not the participants have different perspectives on the fulfilment of each item in the instrument of the practicality test. Thus, using the Rasch model in the development of a teaching technique, "Stepping the 5 Stairs A" looks promising and the analyzed instrument can be used in the further development research.

5. Acknowledgments

6. References

Ariffin, S. R., Omar, B., Isa, A., & Sharif, S. (2010). Validity and reliability multiple intelligent item using rasch measurement model. *Procedia Social and Behavioral Sciences*, 729-733.

Chan, S. W., Ismail, Z., & Sumintono, B. (2014). A Rasch model analysis on secondary students' statistical reasoning ability in descriptive statistics. *Procedia - Social and Behavioral Sciences*, 133-139.

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Method Approaches (4th Edition)* . California: SAGE Publications, Inc.

Kean, J., & Reilly, J. (2018, 8 26). *Classical Test Theory*. Retrieved from Classical Test Theory:https://static1.squarespace.com/static/514fd024e4b0d4d5c3e59e38/t/53bc61
38e4b07f64b2496034/1404854584695/
Kean+Reilly+%282014+in+press%29+Classical+Test+Design.pdf

Mead, A. D., & Meade, A. W. (2018, September 6). *CTT and IRT*. Retrieved from Test Construction using CTT and IRT with Unrepresentative Samples: http://www.csun.
edu/~ata20315/psy427/Topic04_CTT_Reliability.pdf

OoN, D. O. (2018, September 6). *PROF OJERINDE'S LECTURE.* Retrieved from AN EVALUATION OF THE COMPARABILITY OF ITEM ANALYSIS RESULTS: Classical Test Theory (CTT) VS Item Response Theory (IRT): https://ui.edu.ng/sites/.../PROF%20OJERINDE'S%20LECTU

Schumacker, R. E. (2018, September 6). *Clasical Test Analysis*. Retrieved from Applied Measurement Associates LLC: http://www.appliedmeasurementassociates.com/
ama/assets/File/CLASSICAL_TEST_ANALYSIS.pdf

Stage, C. (2018, September 6). *Clasical Test Theory or Item Response Theory: The Swedish Experience*. Retrieved from www.cepchile.cl: http://www.sprak.umu.se/digital
Assets/59/59524_em-no-42.pdf

Sumintono, B., & Widhiarso, W. (2013). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial.* Jakarta: Tim Komunikata Publishing House.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan.* Cimahi: Trim Komunikata.