International Journal of Academic Research in Business and Social Sciences
2017, Vol. 7, No. 2
ISSN: 2222-6990

Validity and Reliability Graphing Calculator Circles Topic Test (GCCTT) Item Based On Table of Test Specification 3 Dimension (3D ToTS) Using Rasch Model: A Pilot Study

Mohd Hafiszudin Ab Samad[1], Raja Lailatul Zuraida Raja Maamor Shah[2], Othman Lebar[3]

[1,2]Department of Mathematics, Faculty of Science and Mathematics,
Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia.
[3]Department of Education Programme, Faculty of Education and Human Development,
Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia.

DOI: 10.6007/IJARBSS/v7-i2/2819 URL: http://dx.doi.org/10.6007/IJARBSS/v7-i2/2819

Abstract
Graphing Calculator Circles Topic Test (GCCTT) Item developed based on the Table of Test Specification 3 Dimension (ToTS 3D) to test the skills of students in terms of graphics calculator skills and circles topic content. GCCTT Item consists of 20 items graphing calculator skills and 14 items of circle topic content. A pilot study using a descriptive quantitative approach carried out in 30 of students Form Two in two schools in Negeri Sembilan using simple random sampling method. To ensure quality GCCTT Item, the item was analysed using the Rasch Model for looking the reliability and validity of items. Reliability derived from the reliability index, meanwhile validity from the point measure correlation (PT-MEA Corr.), Outfit Mean Square (MNSQ) and the Z Standardized (ZSTD). The findings show that GCCTT Item has a high reliability value, 0.85. There are no misfit items detected on graphic calculator skills item except one misfit item detected on circle topic content item. After misfit item dropped, 33 GCCTT Items with quality value and PT-MEA Corr. positive can be produced to test the skills of graphic calculator and circles topic content.

Keywords: Graphing Calculator Circles Topic Test (GCCTT) Item, Graphing Calculator Skills, Graphing Calculator, Technology, Test Item

Introduction

Graphing Calculator Circles Topic Test (GCCTT) Item is a test which was developed to test students in terms of graphics calculator skills and circles topic content. GCCTT Item is developed based on the Table of Test Specifications 3 Dimensional (3D ToTS) comprising the dimensions of the topic, cognitive levels and procedural knowledge of Bloom's Taxonomy. GCCTT Item consists of four subtopics (the circumference of circle concept, circle arc, circle area and circle sector area), Lower Order Thinking Skills (LOTS) which remembers and understands, Higher Order Thinking Skills (HOTS) that apply, analyze, evaluate and create (Churches, 2009) and 11 graphing calculator skills of drawing a circle, radius, diameter and arcs, finding the radius,

diameter, area, angle, arc and calculations operations. GCCTT Item contains 20 items that test students in terms of graphing calculator skills (LOTS item) and 14 items that test circles topic content (HOTS item).

GCCTT Items must be of good quality to test the student's skills of graphing calculators and circles topic content. Therefore, the key to the quality of test items depends on the validity and reliability of the item (Bond & Fox, 2015). Validity refers to what would be measured and the relationship between test items developed with content that has been taught (Churches, 2009; Van Blerkom, 2009) while reliability refers to the consistency of test items administered repeatedly to produce the same score or almost the same every time administered (Ary, Jacobs, Razavieh, & Sorensen, 2010; Howard & Henry, 1988; Lembaga Peperiksaan, 2013). Thus, the analysis item using the Rasch Model to ensure GCCTT Item validity and reliability, as well as recognizing student's achievement and the difficulty level of the items on the same scale of measurement (person-item map*)* (Azrillah Abdul Aziz, Mohd Saidfudin Masodi, & Azami Zaharim, 2013; Bond & Fox, 2015; Edelen & Reeve, 2007).

Methodology

The pilot study was conducted by using quantitative descriptive designs which refer to the collection of information and explain what you want to study (Grimes & Schulz, 2002; Lund Research Ltd, 2015; The Association for Educational Communications and Technology, 2001). A sample of 30 Form Two students in two schools in Negeri Sembilan is selected as the pilot students using simple random sampling method. GCCTT Item has been through the review of the expert (Creswell, 2012; Messick, 1989; Wolfe & Smith, 2007) in the field of Mathematics and a graphing calculator. The results of the review expert, GCCTT Item made improvements before administered to the pilot students. After the pilot study carried out, the data of the GCCTT Item are analysed using Winstep software version 3.72.3.

Findings

GCCTT Item consists of 20 items that test graphics calculator skills and 14 items that test circles topic content. The findings discussed in 3 divisions namely reliability, separation and the validity of the item.

Reliability And Separation Item

According to Table 1, the reliability of GCCTTT item is 0.85 approximate to the value of Rasch Model item reliability, 0.86. Meanwhile the value of Cronbach's Alpha is 0.82. This value indicates the reliability of GCCTT Item within the range that can be accepted as a minimum value of Cronbach's Alpha is 0.8 (Azrillah Abdul Aziz et al., 2013; Bond & Fox, 2015). For the separation of the item, the separation value is 2.39 showing GCCTT Item can be accepted and can be categorised into at least two categories of items, namely easy and hard items (Bond & Fox, 2015).

Table 1
The Reliability and Separation Item

	Total Score	Count	Model		Infit		Outfit	
			Measure	Error	MNSQ	ZTD	MNSQ	ZSTD
Mean	11.0	30.0	0.27	0.57				
S.D	7.3	0.0	1.73	0.33				
Max.	26.0	30.0	4.40	1.82				
Min.	0.0	30.0	−2.90	0.41	0.67	−2.0	0.44	−1.9
Real RMSE	0.67	True SD	1.59	Separation	2.39	Item Reliability		0.85
Model RMSE	0.66	True SD	1.60	Separation	2.43	Item Reliability		0.86

Person Raw Score-To-Measure Correlation = 1.00
Cronbach Alpha (Kr-20) Person Raw Score "Test" Reliability = .82

The Validity Item

The validity of the items measured from the point-measure correlation (PT-MEA Corr.) indicating whether all the items are in one direction with a developed construct. If the item has a negative value of PT-MEA Corr., the item does not measure what should be measured and proper eliminated (Azrillah Abdul Aziz et al., 2013; Bond & Fox, 2015). The positive PT-MEA Corr. showed items have construct validity. Validity refers to the appropriateness of items with the Rasch Model. The misfit item seen from the outfit mean square (MNSQ) exceeds the sum of the mean and standard deviation (outfit MNSQ > Mean + SD) (Azrillah Abdul Aziz et al., 2013) and the Z-Standardized (ZSTD) exceeding 2 (ZSTD > 2) (Bond & Fox, 2015; Linarce, 2012).

From Table 2, the outfit MNSQ in the range of 0.34 to 1.69 while ZSTD value in the range of 0.6 to 1.1. If the value of outfit MNSQ> 1.4 and ZSTD> 2, items considered misfit and proper dropped. However, the findings show no misfit item detected for graphing calculator skill items. Based on the positive PT-MEA Corr. value, this indicates that the item is in one direction with a developed constructs that can measure the skills of graphing calculators.

Table 2
Maximum and Minimum Ranking of Graphing Calculator Skills Item

Ranking	Entry Number	Infit		Outfit		PT-MEA Corr.	Item
		MNSQ	ZSTD	MNSQ	ZSTD		
Maximum	14	1.02	0.2	1.69	1.1	0.37	S3K4
Minimum	6	0.65	-1.0	0.34	-0.6	0.58	S2K2
	Mean			0.96			
	S.D			0.34			

From Table 3, the outfit MNSQ in the range of 0.53 to 2.99 while ZSTD value in the range of −0.6 to 2.2. The findings show item 3 (S1KO3) is a misfit item because the value outfit MNSQ

> 1.7 and ZSTD > 2. Based on the negative PT-MEA Corr. value (−0.1), clearly shows the item is in opposite direction to construct developed which does not measure the circles topic content. Therefore, item 3 (S1KO3) should be dropped because it does not conform to the Rasch Model.

Table 3
Maximum and Minimum Ranking of Circles Topic Content Item

Ranking	Entry Number	Infit		Outfit		Pt-Mea Corr.	Item
		MNSQ	ZSTD	MNSQ	ZSTD		
Maximum	3	1.26	0.7	**2.99**	**2.2**	**-0.1**	**S1KO3**
Minimum	12	0.77	-0.4	0.53	-0.6	0.45	S4KO3
	Mean			1.04			
	S.D.			0.66			

Conclusion

GCCTT Item is analysed based on the Rasch Model for ensuring the quality and appropriate items. From the analysis, the GCCTT Item which consists of graphic calculator skills items and circles topic content items have high reliability value that indicates the GCCTT Item is stable and consistent. The separation of items exceeding 2 shows the difficulty of the items can be separated between the easy item and the difficult item. 20 items which test the graphic calculator skills is maintained while the 14 items which test circle topic content reduced to 13 items. Therefore, 33 GCCTT Items with high quality can be produced that will not only test the circles topic content but also test the skill of students using graphic calculators.

Acknowledgement

The researchers thank the Universiti Pendidikan Sultan Idris (UPSI) for helping this study through the University Grants Special (Education), Project Code 2016-0110-107-01.

Corresponding Author

Mohd Hafiszudin Ab Samad
Department of Mathematics, Faculty of Science and Mathematics,
Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, Malaysia.
cikguhafiszudin@gmail.com

REFERENCES

Ary, D., Jacobs, L., Razavieh, A., & Sorensen, C. (2010). *Introduction to Research in Education* (8th ed.). Cengage Learning. Retrieved from http://books.google.com/books?hl=en&lr=&id=FqF7n0zGJm0C&oi=fnd&pg=PR7&dq=introduction+to+research+in+education&ots=n6Hovzgo67&sig=FaAAiB_Xb7-s3XiA0zgYNyfDM8A

Azrillah Abdul Aziz, Mohd Saidfudin Masodi & Azami Zaharim (2013). *Asas Model Pengukuran Rasch, Pembentukan Skala & Struktur Pengukuran*. Bangi, Selangor: Penerbit UKM.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model. In Applying the Rasch model: Fundamental Measurement in the Human Sciences* (3rd ed.). New York: Routledge.

Churches, A. (2009). Bloom's Digital Taxonomy. Retrieved February 6, 2015, from http://edorigami.wikispaces.com

Creswell, J. W. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (4th ed.). Pearson Education, Inc.

Edelen, M. O., & Reeve, B. B. (2007). Applying Item Response Theory (IRT) Modeling To Questionnaire Development, Evaluation, And Refinement. *Quality of Life Research*, *16*(SUPPL. 1), 5–18. http://doi.org/10.1007/s11136-007-9198-0

Grimes, D. A., & Schulz, K. F. (2002). Descriptive studies: what they can and cannot do. *The Lancet*, *359*(9301), 145–149. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673602073737

Howard, W., & Henry, B. (1988). *Test Validity*. New Jersey: Lawrence Erlbaum Associates.

Lembaga Peperiksaan (2013). *Pentaksiran Kemahiran Berfikir Aras Tinggi*. Putrajaya, Malaysia: Percetakan Surya Sdn. Bhd.

Linarce, J. M. (2012). Winsteps Tutorial 2. Retrieved December 24, 2016, from http://www.winsteps.com/a/winsteps-tutorial-2.pdf

Lund Research Ltd (2015). Descriptive Research Questions. Retrieved November 3, 2015, from http://dissertation.laerd.com/types?of?quantitative?research?question.php#one

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, *18*(2), 5–11. http://doi.org/10.3102/0013189x018002005

The Association for Educational Communications and Technology (2001). What Is Descriptive Research? Retrieved September 14, 2015, from http://www.aect.org/edtech/ed1/41/41-01.html

Van Blerkom, M. L. (2009). *Measurement and Statistics For Teachers* (1st ed.). Taylor & Francis Group. Retrieved from http://www.ebookstore.tandf.co.uk/html/moreinfo.asp?bookid=536959849

Wolfe, E. W., & Smith, E. V. (2007). Instrument development tools and activities for measure validation using Rasch models: part II--validation activities. *Journal of Applied Measurement*, *8*, 204–234.