Using Rasch Measurement Theory to Improve the Functioning of an Assessment for Trigonometry

Sarah Bansilal
Bansilals@ukzn.ac.za
School of Education,
University of KwaZulu-Natal

ABSTRACT: *It is important for teachers to be experts in the content that they teach, and it will be helpful for professional development programmes to design suitable and fair assessments for teachers. The purpose of this study was to use Rasch analysis to improve the functioning of the items in an instrument designed to assess the trigonometry knowledge of mathematics teachers. The participants in the study were 168 high school mathematics teachers who were enrolled in an in-service programme. Fourteen items were rescored to resolve the disordered thresholds that they exhibited. The rescoring resulted in an improved fit of the instrument. Furthermore, this cohort of teachers struggled with higher-level trigonometry questions, especially those which required shifts between different registers of representations. It suggests that the Department of Education needs to provide support to teachers by offering workshops which focus on building up their knowledge in trigonometry by also developing their representational fluency in trigonometry.*

KEYWORDS: **Rasch analysis; rescoring; trigonometry; mathematics teacher knowledge**
➔ Received 30/05/2023 ➔ Revised manuscript received 30/07/2023 ➔ Published 30/12/2023.

1. Introduction

The mathematics learning outcomes in South Africa have been persistently low, especially in mathematics and science. The results for the Grade 12 mathematics national examination for the years 2015 to 2019 reveal that less than a quarter of the learners passed at the 50% level (South African Department of Basic Education [SADBE], 2019; 2020). Furthermore, the number of learners who passed with 50% or more in the Grade 12 mathematics examination in 2015 made up just 5% of the cohort who started school in 2004 (Bansilal, 2017). These low numbers of passes in mathematics in the country have a constraining effect on the economy by limiting the number of people who can access careers in mathematics and science. There are many reasons for the poor outcomes, one of which is teacher's knowledge of the content they teach. Teachers play an important role in helping students achieve positive outcomes since they play a central role in mediating the content with their learners. The well-known McKinsey report summarised this position by stating that the quality of an education system cannot be exceeded by the quality of the teachers in the system (Barber & Mourshed, 2007).

It is important for professional development programmes to also focus on developing the mathematics content knowledge of practising teachers (James et al., 2015). This study emanated from a professional development programme designed to help teachers improve their understanding and teaching of the mathematics they teach. In this small-scale study, I look at an assessment instrument in trigonometry that was used with in-service mathematics teachers, with the aim of using Rasch analysis to improve the functioning of the items. The research questions that underpin this study are:

1. How can an application of the Rasch Measurement Model to an assessment instrument contribute to an improved scoring rubric?
2. To what extent does the empirical ordering of the items described in terms of the item difficulty locations correspond to the increasing cognitive complexity of the items predicted by the education department assessment taxonomy

This study focuses on trigonometry, which is a branch of mathematics that brings together concepts in algebra, geometry and graphs and has applications in many areas of science and mathematics. It is hoped that this study can add knowledge to the area of teacher knowledge of trigonometry while providing insights about challenges in learning trigonometry generally.

2. Literature review

Many studies focusing on the content knowledge of mathematics teachers. have been influenced by Shulman's description of pedagogical content knowledge, which is described as including:

> *... for the most regularly taught topics in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations - in a word, the ways of representing and formulating the subject that make it comprehensible to others. Since there are no single most powerful forms of representation, the teacher must have at hand a veritable armamentarium of alternative forms of representation (Shulman, 1986, p.9).*

Clearly, finding ways of formulating the concept so that it is comprehensible to others requires a sound understanding of mathematics, and this notion of pedagogic content knowledge underscores how important it is for mathematics teachers to have a deep and connected knowledge of the mathematics they teach. However, many studies suggest that mathematics teachers, particularly in South Africa, struggle with the content they need to teach (Bansilal et al., 2014; Mudaly & Moore-Russo, 2011; Taylor, 2011).

Shulman's (1986) description highlights the need for teachers to have a robust knowledge of different representations of concepts. It is crucial that teachers are able to draw upon various representations of a concept. Moore-Russo and Viglietti (2012) explained that it is through mutual interaction between a person's mental representation and other forms of physical representation that a person comes to understand a concept. People with a sound understanding of a concept should know a variety of representations to choose the most appropriate one for a particular situation to solve problems, illustrate properties, or explore relationships (Moore-Russo & Viglietti, 2012). Working with various representations contributes to greater possibilities for mathematical activities across the representations.

A mathematical representation of an object is linked to an underlying semiotic system, which is characterized by a set of signs, a set of rules that governs the production and use of the signs and an associated meaning structure which is based on the relationships between the signs and objects within the system (Ernest, 2006). Duval (2006) notes that semiotic systems of representation are used to label or identify mathematical objects or communicate and allow one to work on a mathematical object. The author distinguishes two types of transformations of semiotic representations that can occur during any mathematical activity (Duval, 2006). The first type, called *treatments*, involves transformations from one semiotic representation to another within the same system or register (Duval, 2006, p. 110). An example of a treatment is transforming the trigonometric expression sin2x into 2sinx.cosx, which is derived from relationships within the trigonometric register. A second type of transformation is that of a conversion, which involves changing the system of representation while preserving the reference to the same object (Duval, 2006). An example of a conversion is expressing the function y = cos x as a graph on the coordinate system. Duval (2006, p. 107) contends that the 'ability to change from one representation system to another is very often the critical threshold for progress in learning. Duval's contention is that treatments command more attention in mathematics, whilst it is conversions which cause the greatest difficulties in mathematics. He argues that conversions come in only because we need to choose 'the register in which the necessary treatments can be carried out most economically or most powerfully'. Another reason put forward for the use of conversions is that they provide *"a second register to serve as a support or guide for the treatments being carried out in another register"* (Duval, 2006, p. 127).

Many problems in trigonometry require the coordination of more than one representational system, and often, a solution cannot be reached if one is confined to activities that are exclusive to one system only (Ubah & Bansilal, 2019). Although the move to a different representation is not essential in all situations, a shift between different representations is often crucial to arrive at a solution to a problem. Duval (2006, p.105) argued that an important aspect of mathematical thinking is that of coordinating two registers of representation simultaneously when he commented that *"the characteristic feature of mathematical activity is the simultaneous mobilization of at least two registers of representation with the possibility at any moment of changing from one to another."*

3. Rasch Measurement theory

Rasch measurement theory (RMT) provides a useful framework for analyzing data from educational, psychological and medical assessments. Central to RMT is the idea that the probability of a test-taker answering an item correctly depends on both the difficulty of the item and the proficiency of the test-taker in the construct being assessed. One of the advantages of using a Rasch analysis is that it computes the proficiency of the test-taker and the item difficulty on one scale.

The Rasch simple logistic model (SLM) for dichotomous items is given in Equation 1 (Andrich & Marais, 2012). In RMT the equation which relates the ability of learners and the difficulty of items is given by the logistic function:

$$P\{X_{vi} = 1\} = \frac{e^{\beta_v - \delta_i}}{1 + e^{\beta_v - \delta_i}} \quad [1]$$

This function expresses the probability of a person *v*, with ability β_v responding successfully on a dichotomous item *I,* with two ordered categories, designated as 0 and 1. Here *P* is the probability of a correct answer; *Xvi* is the item score variable allocated to a response of person *v*, on dichotomous item *i*; β_v is the ability of person v and δ_i is the difficulty of item *i*.

The item characteristic curve (ICC) depicted in Figure 1 shows the alignment of item difficulty and person proficiency. Teachers are represented on the horizontal axis from low proficiency (to the left, towards -5) to high proficiency (to the right towards +3). The expected value for correct response is represented by the vertical axis from 0 to 1. The item is located at a difficulty level of -0.394 logits.

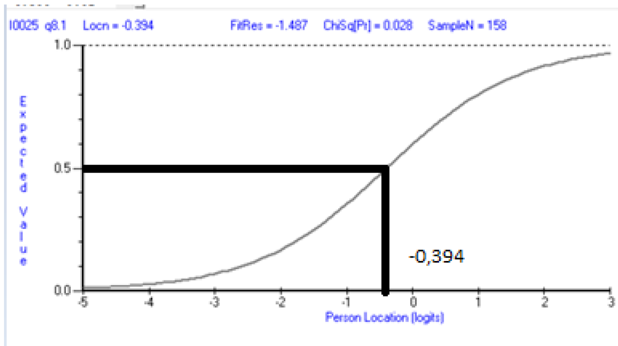


Figure 1. Item characteristic curve for an item located at -0.394 logits.

The corresponding ICC in Figure 1 shows that a person located at the proficiency location of the item difficulty (-0.394) has a probability of 0.5 (expected value of 0.5) of getting the item correct.

A Rasch analysis also generates category probability curves for each item, as shown in Figure 2 for the same item appearing in Figure 1. The curve corresponding to the category of 1 is identical to that of the ICC curve (Figure 1). In addition, the figure also has a curve which denotes the probability of a person scoring 0. The location of the item is identified as the point on the ability scale where the probability curves 0 and 1 intersect. At this point, the probability of a response of either 0 or 1 is equally likely. Because it is a dichotomous item, the person who is located at -0.394 logits (item difficulty) has a probability of 0.5 of either response. The probability of a correct response decreases as proficiency decreases and increases as proficiency increases around this point.

Equation 1 is suitable for items which are dichotomous in nature and has been extended, as shown in Equation 2, to represent the situation when items may be polytomous. Since the response data in this study is polytomous, we use the Partial Credit Model (PCM), which analyses the responses recorded in two or more ordered categories. The equation of the model is expressed in Equation 2 (Andrich, 1978).

$$P\{X_{vi} = 1\} = \frac{e^{(x(\beta_v - \delta_i) - \sum_{k=1}^{x} \tau_{ki})}}{\sum_{x=0}^{mi} e^{(x(\beta_v - \delta_i) - \sum_{k=1}^{x} \tau_{ki})}} \quad \text{Equation 2}$$

Which expresses the probability of a person of ability β_v being classified in a category *x* in a test item of difficulty δ_i, with *m*+1 ordered categories

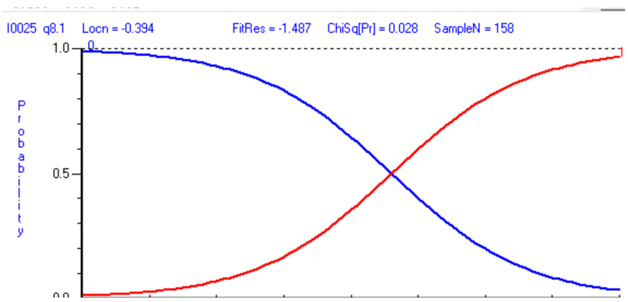


Figure 2. Category probability curve for an item located at -0.394

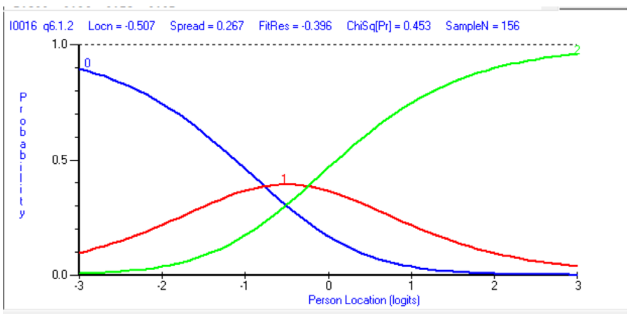


Figure 3. Category probability curve for an item with two thresholds

where $x \in \{1,2, \dots, m\}$ and are the thresholds. The term threshold defines the transition between two adjacent categories, for example, between scoring 0 and 1 (τ_1), or scoring between 1 and 2 (τ_2). This is illustrated below in Figure 3, which shows the category curves for Item 6.1.2.

In Figure 3, there are three category curves corresponding to the probabilities of obtaining a score of 0, 1 or 2. The thresholds, and the categories they define, are naturally ordered in the sense that the threshold defining the two higher categories of achievement is of greater difficulty than the threshold defining the two lower categories of achievement, The first threshold (τ_1), which represents the point where a score of 1 becomes more likely than a score of 0, is about -0.8 logits. The second threshold, where a score of 2 becomes more likely than a score of 1, is approximately -0.2 logits. These thresholds show that progressively more ability is required to score a 0, 1 or 2, respectively, on this item.

4. Methodology

The participants in the study were part of a group of high school mathematics teachers who were enrolled in an in-service programme. There were 168 participants whose records were captured for this study. The items in the data collection instrument were drawn from a previous Grade 12 examination paper usually taken by school students. In this article, we focus on 16 items based on trigonometry to see how well those items worked as an assessment of proficiency in trigonometry as a whole.

When data is analyzed, it is common practice to discuss how well the model fits the data. However, with respect to RMT, the requirement is that the data fit the model in order to claim measurement within the models' framework. The fit statistics are used to help detect when the data does not fit the model as expected, and this allows us to diagnose some reasons for the misfit.

The scores for the teachers in each of the items were recorded in Excel. The data were cleaned and then exported into RUMM 2030, where further analysis was undertaken. The Department of Education utilizes an assessment taxonomy to guide the design of assessments in mathematics. The taxonomy distinguishes between four cognitive levels of the items according to the Department of Basic Education (DoBE) assessment taxonomy. Table 1 presents a summary of these levels.

Each of the items was categorized into the four cognitive levels by three experts individually. The experts then discussed the categories until agreement about the item levels were agreed upon. Note that the classification of the items according to these taxonomy levels is not cast in stone and may be interpreted slightly differently by different people depending on the cohort of learners and the background. The details of the items with the original fit statistics, the location estimates, as the categorization according to the assessment taxonomy level of Table 1 are presented in Table 2.

5. Results

From the initial Rasch analysis, the summary statistics (Table 3), person-item location distribution (Figure 4) and person-item threshold distribution (Figure 5) were generated. Table 2 presents the initial summary statistics, which shows the item mean as 0 (as set by the model)

Table 1. Descriptors of each level of the assessment taxonomy used by the Department of Basic Education

Cognitive levels	Description of skills to be demonstrated
(1)Knowledge	Straight recall Identification of the correct formula on the information sheet (no changing of the subject) Use of mathematical facts Appropriate use of mathematical vocabulary
(2)Routine Procedures	Estimation and appropriate rounding of numbers Proofs of prescribed theorems and derivation of formulae Identification and direct use of correct formula on the information sheet (no changing of the subject) Perform well-known procedures Simple applications and calculations which might involve a few steps Derivation from given information may be involved Identification and use (after changing the subject) of the correct formula Generally similar to those encountered in class
(3)Complex Procedures	Problems involve complex calculations and/or higher-order reasoning There is often not an obvious route to the solution Problems need not be based on a real-world context Could involve making significant connections between different representations Require conceptual understanding
(4)Problem Solving	Non-routine problems (which are not necessarily difficult) Higher-order reasoning and processes are involved Might require the ability to break the problem down into its constituent parts

Table 2. Brief descriptions of items with original fit residual statistics, item location and cognitive levels

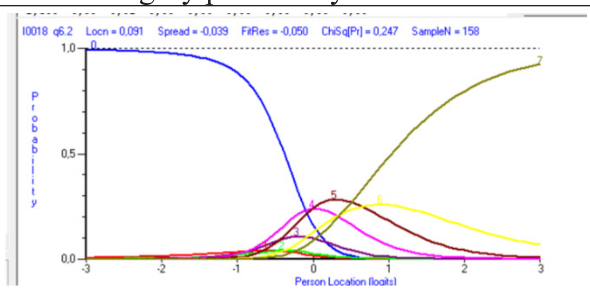
Item	Item with comments	(FR)	Locn	Cog lev
15	6.1.1 Identifying and using the correct formula to express trig ratio of cos	-0,544	-0,363	2
16	6.1.2 Identifying and using the correct formula for double-angle	-0,948	-0,218	2
17	6.1.3 Rewriting 4° as (32°- 28°) & applying correct compound angle formula	-0,798	0,128	3
18	6.2 Using the correct double angle formula & finding a general soln. of a trig. eqn	-0,05	0,091	3
19	6.3.1 Identifying values of x for which the given trig expression undefined	0,325	0,569	3
20	6.3.2 Proving an identity by identifying & using a correct double-angle formula.	0,876	-0,135	3
21	7.1 Identifying and using the cosine rule using the correct triangle	-0,627	-0,558	2
22	7.2 Identifying and using the sine rule to get the correct angle in the correct triangle	-0,01	-0,18	2
23	7.3 Using area formula correctly utilizing result in 7.1	-0,253	-0,335	1
24	7.4 Identifying that EF is EG + GF first & then using correct trig ratio for EG; etc	1,93	-0,334	2
25	8.1 Identifying period from the sketched graph -	-1,487	-0,394	1
26	8.2 Working out the amplitude of the new function described in terms of the sketched function	0,894	-0,372	2

Item	Item with comments	(FR)	Locn	Cog lev
27	8.3 Sketching graph that has been translated 30 to the right + correct endpoints	-0,603	-0,239	2
28	8.4 Identifying the no. of intersection point for the two drawn graphs above	-2,019	0,509	2
29	8.5 Identifying the region for which the graph of g in 8.2 is on/above the X-axis.	-1,762	0,561	3
30	Identifying regions for which f'(x) > 0 and g'(x) > 0	-1,047	1,271	4

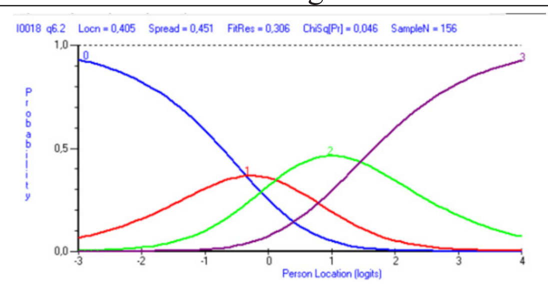
Table 3. Initial summary statistics

	ITEMS [N=16]		PERSONS [N=168]	
	Location	Fit residual	Location	Fit residual
Mean	0.0000	-0.3827	-0.0800	-0.1050
SD	0.4937	1.0392	1.0157	0.6997
			Person separation index 0.85186	

Initial category probability curves for Item 18



After rescoring Item 18



Question with marking memorandum

Determine the general solution of: $\sin^2 x + \cos 2x - \cos x = 0$ (7)

Solution	Marks
$\sin^2 x + \cos 2x - \cos x = 0$	
$\sin^2 x + (\cos^2 x - \sin^2 x) - \cos x = 0$	✓ $(\cos^2 x - \sin^2 x)$
$\cos^2 x - \cos x = 0$	✓ $\cos^2 x - \cos x = 0$
$\cos x(\cos x - 1) = 0$	✓ factors
$\cos x = 0 \ \ or \ \cos x = 1$	✓ $\cos x = 0 \ \ or \ \cos x = 1$
$x = \pm 90° + k.360° \quad or \ x = 0° + k.360° \quad k \in Z$	✓ $90° + k.360°$
$= k.360°$	✓ $k.360°$
	✓ $x = -90° + k.360°$
(i.e. $x = 90° + k.180°$ or $x = k.360° \pm 90°, k \in Z$)	(7)

Figure 4. Rescoring of an item with disordered thresholds

and the person mean as- 0.0800. The standard deviation for the item location is 1.0392, which is just above the ideal value of 1, while the standard deviation of the person location is 1.0157, which is just above 1.

The mean of the item fit residual (-0.3827) is close to zero. The standard deviation of the item fit residual is 1.0392, close to 1, showing that the fit does not vary more than expected. The mean of the person fit residual is approximately

-0.1050, showing that it is close to zero. The standard deviation of the person fit residual is 0.6997, which is slightly smaller than 1, showing that the distribution of the person fit residuals is slightly more clustered than the ideal situation.

The person separation index from Table 3 is over 0.85. This shows that the estimation of the person's ability is consistent across the model. Based on this separation figure, one calculates the reliability as 0.683, which suggests that the test did not separate the persons as well as it could.

In RUMM 2030, item fit z-score transformed residuals between ±2.5 are deemed adequate fit to the model (Pallant &Tennant, 2007). The initial analysis identified no items as having RUMM 2030 misfit statistics outside these recommended limits, as shown in Table 2. Each item was analyzed for the content and marking allocation, while the Item characteristic curves (ICC) and category probability curves (CPC) were also examined. Based on these analyses, the items were rescored post-hoc, where necessary.

The Item characteristic curves (ICC) and category probability curves (CPC) for each item were also examined, and based on these, the items were rescored if such a rescoring was supported by the qualitative analysis. If a rescoring was suggested by the analysis and the rescoring was supported by the qualitative analysis, then the item was rescored. This rescoring process resulted in improved fit residual statistics overall. Figure 4 demonstrates an example of an item which initially had disordered thresholds and which improved after the rescoring.

The initial category curves in the first column of the first row in Figure 4 show disordered thresholds, which may signal that certain categories are not working well (Andrich, 2005). The categories corresponding to scores of 1, 2, 3, and 6 do not seem to be functioning as intended. At no point on the horizontal axis is a score of 1, 2, 3 or 6 most likely. In such a situation (Andrich, 2005) suggests a rescoring to try to eliminate the disordering of the thresholds. However, note that not all authors see the disorder of the estimated parameters as an indication of a misfit of data (Adams et al., 2012). In this situation, I followed

Table 4. Rescoring details for Item 18

Original score	Revised score
0	0
1	0
2	1
3	1
4	1
5	2
6	2
7	3

the advice of Andrich (2005) in trying to resolve the disordered thresholds by first studying the marking rubric and checking if the mark allocation made sense. I first conducted a content analysis of the item and analyzed the rubric. I then rescored the item with a maximum score of 3, as shown in Table 4.

As seen in the right-hand side of the second row of Figure 4, the rescoring resulted in ordered thresholds, showing that a higher mark required greater proficiency than a lower mark. By rescoring the item, the thresholds are able to contribute consistently to a scale to measure the construct of interest, which is teachers' proficiency in trigonometry. In the RHS of Figure 1, the thresholds and the categories they define are naturally ordered in the sense that the threshold defining the two higher categories of achievement is of greater difficulty than the threshold defining the two lower categories of achievement. The first threshold (τ_1), which represents the point where a score of 1 becomes more likely than a score of 0, is about -1.1 logits. The second threshold, where a score of 2 becomes more likely than a score of 1, is approximately 0.2 logits, whereas the third threshold is approximately 1.5. These thresholds show that progressively more ability is required to score 0, 1, 2 or 3 marks, respectively, on this item (Van Wyke & Andrich, 2006, pp. 13-14)

There were 14 polytomous items which indicated discorded thresholds, and these were also rescored based on a consideration of the scoring rubric. The rescoring resulted in an improved fit to the Rasch model, as the Person

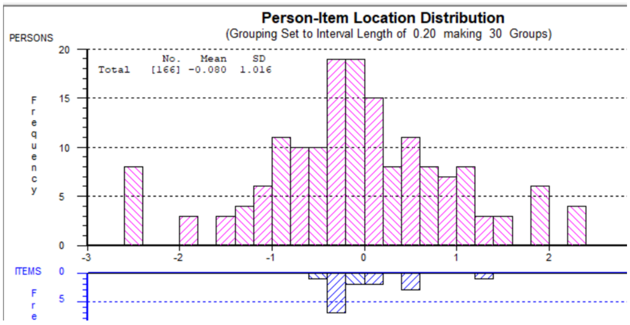


Figure 5. Original Person- item location distribution

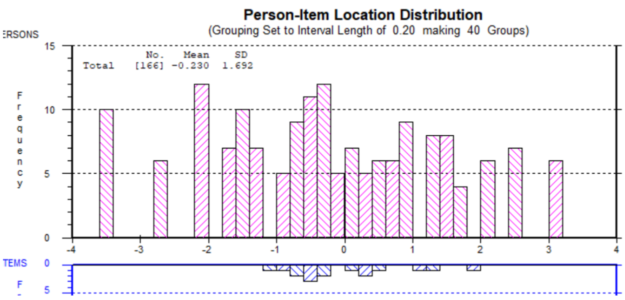


Figure 6. Final person-item location distribution after rescoring

separation index improved from 0.85186 to 0.86060, as shown in Table 5 below. None of the items displayed misfit.

In terms of the distribution of the persons by ability, the distribution presented in Figure 6 indicates a better balance than the original distribution represented in Figure 5.

The items ranged from-0.558 to 1.271 and after rescoring it was -1.060 to1.920, providing a greater spread. We can see that the SD for person locations has increased in the person-item location distribution from 1.01857 to 1.6918. The person locations ranged between – 2.4 and

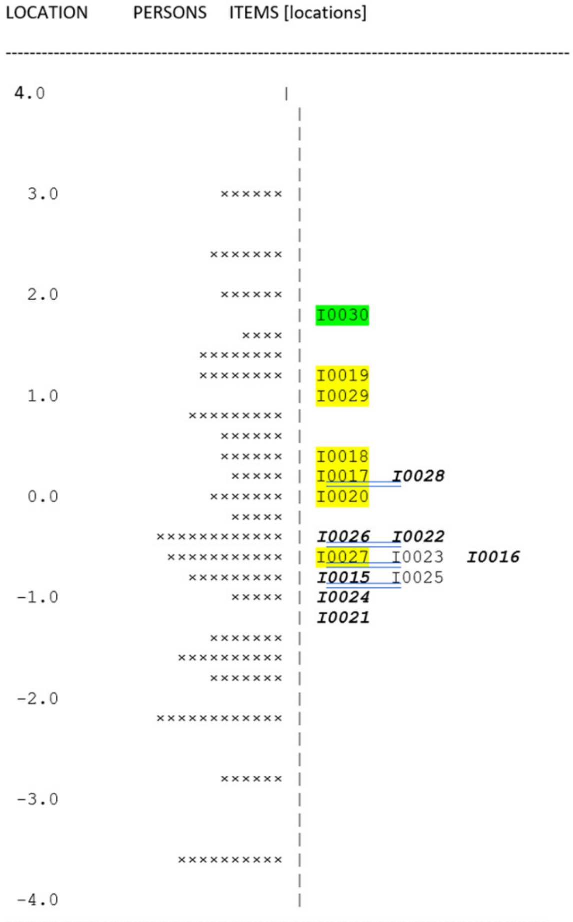


Figure 7. Wright map showing item and person location

Key: L1- normal font; L2- bold+ italics; L3- yellow highlights; L4- green highlights

2.2 logits and after rescoring the spread widened to between – 3.53 to 3.05, thus providing greater discrimination of person proficiencies. In terms of the threshold distribution, the number of thresholds decreased from 47 to 23. By improving

Table 5. Initial and final summary statistics

	ITEMS [N=16]		PERSONS [N=168]	
	Location	**Fit residual**	**Location**	**Fit residual**
Mean	0.0000 (0.000)	-0.3827(-0,3045)	-0.0800(-0.2300)	-0.1050(-0.1943)
SD	0.4937 (0.8643)	1.0392 (1.0723)	1.0157(1.6918)	0.6997(0.7783)
			Person separation index 0.85186 (0.8606)	

Draw the graph of $g(x) = \cos(x - 30°)$ for $x \in [-180° ; 180°]$ on the diagram

Figure 8. Instruction for Item 17

the scoring rubric, there is a greater spread of the person difficulties; the test has still not separated the persons very well since there are many test takers who were not able to be 'measured' by the items, and since there are only 16 items, it is unlikely to produce good person separation even if they target the population better.

In terms of the empirical ordering of the items, there was a match between the predicted ordering according to the theoretical assessment taxonomy and the actual empirical ordering based on the Rasch analysis, with the items predicted at higher cognitive levels being generally higher than those at lower cognitive levels. However, not all items followed this trend.

Item 27 was classified as a Level 3 question, but empirically, the teachers found it easier than the Level 2 problems. It appears in Figure 8.

This may be because graph sketching is carried out mainly by using the calculator, so there is no need for the reasoning about the shifting of graphs that was expected. Another unexpected result was that of item 28, which was classified as a Level 2 item but empirically was experienced as more difficult than the other Level 2 items. It may be that the language was a problem, and people did not realize they were asking for the number of intersection points and not the intersection points itself.

In looking at the Wright map, one might discern three levels of items, as shown in Figure 10.

From Figure 6, we can discern three clusters of difficulty. The first one comprises items 30, 19 and 29, which were level 3 and 4 items. Items 30 and 29 were based on interpretation of the trigonometric graphs, which required teachers to coordinate two or more different representations. As noted by Duval, mathematical activity often requires the simultaneous mobilization of more than one register of representation. For example, given two trig graphs, they needed to work out Item 30, shown in Figure 11.

Here, they needed to work with the graphical representation of two graphs while coordinating

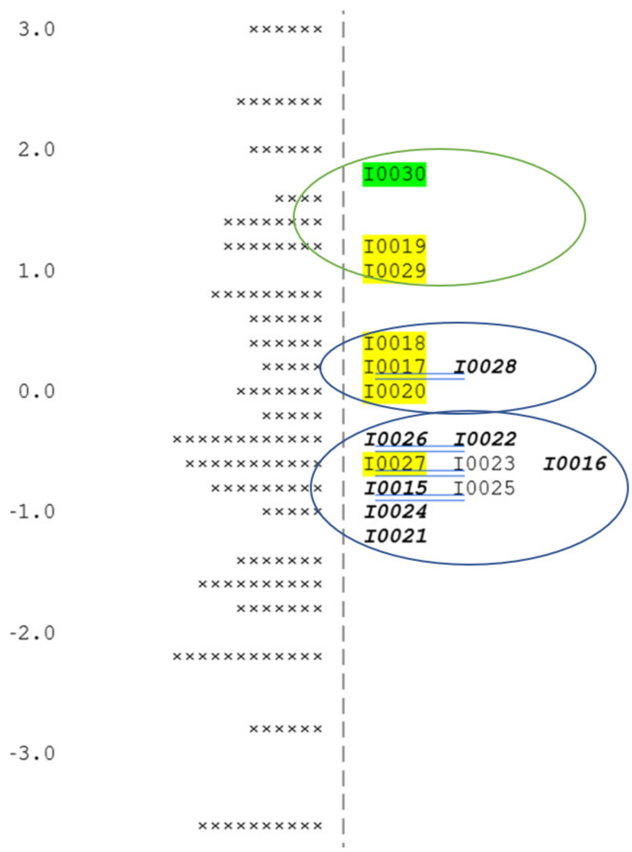


Figure 10. Wright map with three levels of items

Use the graph to determine the number of solutions for $-2\sin x = \cos(x - 30°)$ $x \in [-180° \; ; \; 180°]$.

Figure 9. Instruction for Item 28

Prove that $\dfrac{\cos 2x . \tan x}{\sin^2 x} = \dfrac{\cos x}{\sin x} - \tan x$ for all other values of x.

Figure 12. Item 20

For which values of x is $f'(x) < 0$ and $g'(x) > 0$?

Figure 11. Item 30

the algebraic representation, asking about the derivatives of the two graphs and simultaneously work out the intervals over which the two conditions were satisfied. This study shows that working simultaneously with these different representations registers was experienced as very difficult by the teachers. The results from the Grade 12 2021 mathematics examinations showed similarly that the items based on the interpretation of graphs were the most difficult for learners (SADBE 2020).

The second cluster of items are those which required complex manipulation of trigonometric expressions based on algebraic rules and trigonometric relationships. These require coordination between the trigonometric and algebraic registers, for example, Item 20 below:

Item 20 required the identification of the most suitable identity for cos 2x, followed by a simplification of the trigonometric expressions using the algebraic register and treating the trigonometric expressions as variables.

The third cluster of items were the ones that were experienced as easiest, and these were mainly a combination of level 1 and 2 items. Generally, these items were based on simple manipulation of trig expressions, calculations and recall of formulae, e.g. "What is the period of the graph in the diagram?".

In terms of this cohort of teachers who took this test, there were few teachers who were able to cope with the items in the highest cluster. The difficulty estimate for Item 30 was 1.920, and there were only 19 teachers (11%) whose proficiency levels were higher than these, which means that they were easily able to cope with these items. This suggests that the teachers needed more opportunities to work with such trigonometric problems, which they will need to teach to their students.

Of concern is the large number of teachers whose proficiency levels were lower than all of the items. There were 52 teachers who comprised 30% of the group whose person location was lower than the item difficulty estimates of all the items. This confirms that trigonometry is experienced as a difficult section for learners and their teachers.

5. Conclusions

The purpose of this study was to use Rasch analysis to improve the functioning of the items in a trigonometry assessment done by teachers. Many of the items had disordered thresholds, meaning that the scores did not contribute consistently to a scale that measures their trigonometry proficiency. These items were rescored to resolve disordered thresholds. The rescoring resulted in an improved fit of the instrument. The study is significant because it shows the effect of redundant marks on the ordering of items according to difficulty. It may seem logical to allocate more marks for an item that is considered as difficult. However, when marks are redundant, the allocation of marks to the items in the measuring instrument does not communicate the proficiency of learners in a fair way. The marks do not contribute consistently to a scale in which we have confidence that it is able to represent the proficiency of students, where a higher location means that a person has a greater proficiency than a person on a lower level. As seen in this study, by improving the scoring rubric, the test has been able to distinguish better between items of different difficulty than the original one, thus providing a greater degree of precision. It is important to note, however, that the person separation was not very good, and we would need more items, especially at the upper and lower bounds of person achievement. This test has been able to identify the group of people at the bottom of the distribution, which is crucial to provide them with additional support in this area since they are struggling with most of the concepts.

The empirical ordering of the items, according to the difficulty of the Rasch model, generally supported the levels of difficulty according to the education department's assessment taxonomy, where items categorized at lower cognitive levels were easier than those at higher cognitive levels.

The study has identified that teachers may also struggle with the trigonometric concepts that they teach, which is a concern. It suggests that the education department needs to support teachers by offering workshops and professional development opportunities to practising teachers. This cohort of teachers struggled with higher-level trigonometry questions, especially those which required shifts between different registers of representations. As teachers who will be mediating this content with their students, it is important for them to be given further support in the content that they teach. Teachers need access to professional development programmes that can improve their trigonometry knowledge.

Acknowledgements

Funding for this study was given by the National Research Foundation (South Africa), grant no. 129308.

References

Adams, R. J., Wu, M. L., & Wilson, M. (2012). The Rasch rating model and the disordered threshold controversy. *Educational and Psychological Measurement*, 72(4), 547-573.

Andrich, D. (2005). *The Rasch model explained*. In S. Alagumalai, D. D. Durtis, & N. Hungi (Eds.), Applied Rasch measurement: A book of exemplars (pp. 308-328). Berlin, Germany: Springer-Kluwer

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.

Andrich, D & Marais, I. (2012). Introduction to Rasch measurement of Modern Test theory. Course Notes: University of Western Australia.

Barber M & Mourshed M (2007). How the world's bestperforming school systems come out on top. London, England: McKinsey. Available at https://www.mckinsey.com/~/media/mckinsey/indu stries/public%20and%20social%20sector/our%20insights/how%20the%20worlds%20best%20performing%20school%20systems%20come%20out%20on%20top/how_the_world_s_bestperforming_school_systems_come_out_on_top.pdf . Accessed 12 September 2020.

Bansilal S (2017). The difficulty level of a national assessment of Grade 9 mathematics: The case of five schools. *South African Journal of Childhood Education*, 7(1):a412. https://doi.org/10.4102/sajce.v7i1.412

Bansilal, S., Brijlall, D., & Mkhwanazi, T.W.(2014). An exploration of the common content knowledge of high school mathematics teachers. *Perspectives in Education*, *32*(1), 34-50.

Duval R (2006). A cognitive analysis of problems of comprehension in the learning of mathematics. *Educational Studies in Mathematics*, 61, 103-131.

Ernest P (2006). A semiotic perspective of mathematical activity: The case of number. *Educational Studies in Mathematics*, 61, 67-101.

James, A., Bansilal, S., Webb, L., Goba, B., & Khuzwayo, H. (2015). Teacher professional development programmes in MST for developing contexts, Africa Education Review, 12(2), 145-160, Doi: 10.1080/18146627.2015.1107970

Moore-Russo, D., & Viglietti, J. (2012). Using the K5 Connected Cognition Diagram to analyze teachers' communication and understanding of regions in three-dimensional space. *Journal of Mathematical Behavior*, *31*(2), 235−251.

Mudaly, V., & Moore-Russo, D. (2011). South African teachers' conceptualizations of gradient: a study of historically disadvantaged teachers in an Advanced Certificate in Education programme. *Pythagoras*, *32*(1), 1-8.

Robinson, M., Johnson, A.M., Walton, D.M. *et al.* A comparison of the polytomous Rasch analysis output of RUMM2030 and R (ltm/eRm/TAM/lordif). *BMC Med Res Methodol* **19**, 36 (2019). https://doi.org/10.1186/s12874-019-0680-5

South African Department of Basic Education, Republic of South Africa 2020. Report on the 2019 National Senior Certificate (Diagnostic Report, Part 1). Pretoria: Author. Available at https://www.education.gov.za/Portals/0/Documents /Reports/DBE%20Diagnostic%20Report%202019 %20-%20Book%201.pdf?ver=2020-02-06-140204- 000. Accessed 31 December 2022.

South African Department of Basic Education, Republic of South Africa 2019. National Senior Certificate 2018 Diagnostic Report Part 1. Pretoria: Author. Available at https://www.education.gov.za/Portals/0/Documents /Reports/FINAL%20DIAGNOSTIC%20REPORT %20NON%20LANGUAGES%202018.pdf?ver=20 19-01-24-125928-000. Accessed 31 December 2022.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational researcher*, *15*(2), 4-14.

Taylor N (2011). The national school effectiveness study (NSES). Johannesburg, South Africa: Joint Education Trust. Available at https://www.jet.org.za/resources/summary-forsynthesis-report-ed.pdf. Accessed 31 December 2022

Pallant, J.F., &Tennant, A.(2007). An introduction to the Rasch measurement model: an example using the hospital anxiety and depression scale (HADS). Br J Clin Psychol 2007;46 Pt 1:1-18. Doi:https://doi.org/10.1348/014466506X96931.

Ubah, I & Bansilal, S. (2019). The use of semiotic representations in reasoning about similar triangles in Euclidean geometry, *Pythagoras*, *40*(1) https://doi.org/10.4102/pythagoras.v40i1.480

Wright, B. D. (1997). A history of social science measurement. *Educational Measurement: Issues and Practice* (Winter), 33-45.