Psychometric Properties of Matriculation Entrepreneurial Thinking Scale Using Rasch Model

Amal Harun, Norasmah Othman
National University of Malaysia, Malaysia

Abstract
Entrepreneurial thinking is an indicator of individual excellence that is much-needed in facing the challenges of IR 4.0. Currently, there are many psychometric instruments to measure entrepreneurial thinking, however these are not suitable for measuring the thinking of entrepreneurs among matriculation students. Thus, Matriculation Entrepreneurial Thinking Scale (METS) was developed as a new psychometric instrument. Therefore, this study focuses on the aspects of METS psychometric testing based on Rasch analysis in the context of matriculation students. A total of 378 matriculation students for the 2020/2021 session from two matriculation colleges were selected using the multistage cluster sampling method. After being analyzed using Winstep 3.71 software based on the Rasch model, 55 out of the 66 METS items had satisfied the psychometric characteristics. Ten items were dropped, but five items which were negative statements were re-entered with corrections to positive statements. The strength of Rasch analysis on calibration scale was also reported to determine the suitability of the four-point Likert scale. It is recommended that further studies should examine the effects of re-inserting negative items after being converted to positive and looking at the GDIF aspects of METS items. Good METS items are expected to benefit the measurement of new traits among future matriculation students.
Keywords: Entrepreneurial Thinking, Matriculation, Rasch, Psychometric, Instrument

Introduction
When it comes to facing up to the challenges of the Industrial Revolution 4.0, the country is in desperate needs of knowledgeable and skilled human capital. The World Economic Forum (2020) had prepared a list of the skills that students need to master in 21st century. Among them are problem solving skills, technology skills, self-management and leadership skills. These skills are seen as relevant in ensuring the marketability of students in the working world later on (Rajadurai et al., 2018). And due to that reason, there is a need to produce entrepreneurial minded students in order to face the challenges of a competitive world (Bacigalupo et al., 2016; Buang et al., 2009).

Initially, the entrepreneurial thinking is understood as the thinking style of an entrepreneur who is free from following certain concepts or theories, and more towards following intuition (Blume & Covin, 2011). The implications of entrepreneurial thinking can be seen through the behaviour and success which they manage to achieve (Hnátek, 2015). It is generally considered that entrepreneurial thinking is only focused on creating entrepreneurs and businesses, however, it stretched more as an important set of skills in human capital development, meeting the demand of job market and facing competition (Bacigalupo et al., 2016). The concept of entrepreneurial thinking can be expanded within the context of an individual life as a way of thinking and behaviour that mimics the thinking of an entrepreneur (Dhliwayo, 2008).

Leffler (2019) characterizes students who are entrepreneurial minded as the ability to think critically, creatively and innovatively. While, Edwards-Schachter et al. (2015) explained that students with entrepreneurial thinking are those who are efficient in solving a problem and wisel in managing existing resources optimally. In other words, the student is thinking creatively. The characteristics mentioned are among the skills that are value-added for a student in addition to academic qualifications (Aliu & Aigbavboa, 2021).

The importance of having this entrepreneurial thinking has raised an urge towards development of entrepreneurial thinking among students. This effort has been implemented through the structuring of education policy when the characteristics of an entrepreneur are applied in education (Buang, Halim & Meerah, 2009). Buang, Halim and Meerah (2009) have proposed an integration process to be done which include science process skills and entrepreneurial thinking also known as entrepreneurial science thinking. This thinking integrates the concepts of problem-solving and entrepreneurial elements in order to train students to search for ideas to solve problem creatively (Syukri et al., 2013). The application of entrepreneurial thinking is also implemented at the educational management level. According to a study by Othman et al. (2006), educational institutions have taken steps towards the formation of entrepreneurial thinkings among teachers and administrators. Thus, individuals who are entrepreneurial thinkinged should be an indicator of personal and organizational excellence.

Recently, the entrepreneurial thinking began to be seen as an indicator of personal and organizational excellence, hence, there are various efforts being done towards measuring the entrepreneurial thinking. Ishak (2014) developed an instrument to measure the readiness of science teachers in integrating the entrepreneurial thinking in schools, while Syukri et al. (2013) developed an instrument to measure teachers' pedagogical knowledge in teaching entrepreneurial science thinking. Entrepreneurship Science Thinking Test for primary school students was already been developed by Ahmad and Abdullah (2020), while the Entrepreneurship Thinking Instrument was developed by Hariyaty (2008). Looking into the details, the existing instruments are difficult to adapt in terms of measuring the thinking of entrepreneurs, especially in the context of students' life in matriculation.

Since, the existing measuring instruments of entrepreneurial thinking are limited, researchers had developed new instruments to measure entrepreneurial thinking specifically for matriculation students. The current existing instruments are found to measure entrepreneurial thinking in the context of school administrators and primary school students

only. Thus, new instruments are developed with the addition of several dimensions as compared to existing instruments. Nevertheless, in order to find out the suitability of the instrument in measuring the entrepreneurial thinking of matriculation students, psychometric characteristics need to be tested before the instrument can be used. Therefore, this article discusses the psychometric characteristics of the entrepreneurial thinking instrument based on the Rasch measurement model.

Literature Review
Entrepreneurial Thinking

Entrepreneurial thinking is the way that individual thinks and acts in order to identify opportunities in obtaining profit or benefit. Entrepreneurial thinking also refers to the cognitive aspects which are critical thinking, creativity and innovation. Entrepreneurial minded individuals are perceived to be able to solve complex problems and sensitive to the surrounding environment (Mohamad et al., 2019). Entrepreneurial minded individuals are able to make decisions (Winnaar & Scholtz, 2018) and are willing to take risks (Lumpkin & Dess, 1996). Holm et al. (2013) explained that these individuals are able to face uncertain situations by converting opportunities for personal or community benefit through their creativity, in addition to using their potentials and skills.

Among the dimensions of the entrepreneurial thinking are creativity and innovation (Ishak et al., 2017), critical thinking (Hancock et al., 2020), smart in sensing opportunities (Clausen 2020; Ishak, 2013), daring to take risks (Akbay & Delibalta; 2020), and ability to work in teams (Anwar & Menekse, 2020; Nelson, 2018). Critical thinking skills include decision making skills and problem solving skills (Hancock et al., 2020). Hence, this study adapts Timmons Model (2005) in terms of the context of life as a student who will place himself in the job market.

Timmons Model consists of three main elements in developing a business, namely opportunities, resources and teams. The element of opportunity refers to the ability to identify opportunities and daring to take risks. The second element is the resources referring to creative, innovative and critical thinking skills. Meanwhile, critical thinking skills refer to decision-making skills and problem-solving skills. The third element is teamwork skills. Therefore, all the elements discussed in this model are used as dimensions in measuring the entrepreneurial thinking.

Rasch Model

In 1960, Georg Rasch who was a mathematician had introduced the Rasch model which is also known as the one-parameter model. The parameter particularly refers to the difficulty parameter only. The Rasch model is based on the Item Response Theory (IRT) or known as Latent Trait Theory. The Rasch model uses IRT to analyze the data in regards to the validation of constructed psychometric instruments. This model assumes that an individual's response to an item is influenced only by the difficulty of the item and the individual's ability to answer the item (Bond & Fox, 2007).

The advantage of Rasch analysis is being able to meet the measurement needs. It can convert raw scores that are often influenced by items and samples towards linear measurements with similar interval values (Wright & Stone, 2004). This shows that the Rasch model is able to convert ordinal scales to interval scales with logit units. This article explains the psychometric features of METS based on some of the assumptions of the Rasch model that must be adhered to, namely item fit, unidimensionality, polarity, local independence,

and scale calibration. Aspects of reliability are also reported through the reliability index and the separation index.

Unidimensionality

Unidimensionality refers to the condition where the items in an instrument only measure one dimension (Wright & Master, 1982). Unidimensionality is explained through the Principal Component Analysis (PCA) procedure with four criteria that must be followed. First, the value of raw variance explained by measures should reach a minimum of 40% in order to show that the instrument has good unidimensionality (Ariffin et al. 2013). Second, the value of unexplained variance in the first contrast should not exceed 15% (Fisher, 2007). The third criterion describes the Eigen value which is less than 5.0 (Linacre, 2005), explain the absence of a second dimension. The fourth criterion is the ratio of variance explained by measure to unexplained variance in the first contrast should be at a minimum rate of 3:1 (Linacre, 2018).

Item Fit

Item fit refers to the extent of which the data corresponds to the Rasch model (Ariffin et al., 2010). The item developer can identify the item fit based on the MNSQ infit and outfit values. Theoretically, MNSQ is the ratio of whether an observation meets the expectations. Thus, the ideal value of MNSQ is 1. Bond & Fox (2015) explains that the values which are less than 1.0 indicate that the observations are too predictable due to redundancy or data is too high on the model. Values greater than 1.0 indicate unpredictability due to noise or data does not fit the model well. The MNSQ value of the item must be in the range of 0.77 - 1.30 logits (Fisher, 2007). While Bond and Fox (2015) emphasized that the appropriate range for the Likert scale is between 0.6 logits to 1.4 logits.

Item Polarity

Item polarity is generally used to check that all items are aligned in the same direction on the latent variable (Bond & Fox, 2007). Positive values are indication that all items in the instrument function in one direction in terms of measuring a particular instrument construct. Negative value index usually indicates that an item or individual responds in conflicting manner to the variable (Linacre, 2005). Bond and Fox (2007) highlighted that the Point Measesure Correlation (PTMEA Corr) value must be positive and 0.3 and above. Items with good polarity should be in the range of 0.3 to 0.6 logits. While, negative PTMEA Corr value suggest that it did not measure what should be measured, thus this item should be dropped (Aziz et al., 2013).

Local Independence

Local independence explains the extent to which an individual's level of ability responses towards any item is unrelated to other items within the same construct. According to Linacre (2012), a local freedom range that is a correlation value of less than 0.70 indicates that the items in the construct are not interrelated with each other. Balsamo et al. (2014) and Gibbons et al. (2011) set a correlation value of less than 0.3 as good. Although the correlation exists, the strength of the correlation is weak.

Scale Calibration

Scale calibration is a unique testing found in the Rasch model which is used to identify the effectiveness of scale options used in instruments. Discussions on scaling reviews were

based on the criteria outlined by Linacre (2002) and Bond & Fox (2015). This consideration was made to ensure that the four-point scale used is appropriate for METS. The four-point scale used were; (1) strongly disagree, (2) disagree, (3) agree and (4) strongly agree. There are six criteria that must be followed in the process of scale calibration (Linacre, 2002). The first criterion describes the number of observations must be more than 10 for each category. Higher observation values indicate better expectations for the accuracy of a score. Second, the shape of the curve for each category must display a clear peak. Third, the average measurement value of observations for each category increases in line with the category scale. Consistent improvement indicates a normal and uniform response pattern.

Fourth, the value of the MNSQ outfit must be less than 2.0 logits. Meanwhile, scores exceeding 2.0 logits indicate the degree of interference for the unexplained variance. Fifth, the threshold value must increase in line with the increase in category. Sixth, the restriction category must exceed the value of one, but not exceeding the value of five for a four-point Likert scale (Linacre, 2002). Rasch's analysis will help to verify the probability of the response as being evenly distributed. The calibration structure is evaluated to confirm whether the classification level is applicable. The 's' separation value should be in the range of 1.4 <s <5.0. If the value of s is less than 1.4, then the rating will be combined. On the other hand, if the value of s is more than 5.0, then the rating will be separated.

Methodology

This quantitative study used an online survey design. Google Form was used as the study was conducted in the Covid19 pandemic environment. The researcher used group sampling in stages. The first stage was the random selection of colleges by zone, while the second stage involved the selection of one college for each zone. A total of 378 students had responded from the two selected colleges. The number of samples is sufficient based on the calculation of Raosoft Sample Size Calculator (Raosoft, 2004) with 5% error and 95% confidence level. Students from the two colleges that were involved had similar characteristics to students of other matriculation colleges in terms of age and eligibility for admission to the matriculation program. A large number of students revealed that the participation of students came from various majors and races.

The research instrument used for this research was METS questionnaire which contains six dimensions namely; (1) decision making skills (items D1-D8); (2) problem solving skills (items P9-P14); (3) risk-taking courage (items R15-R29); (4) creative and innovative thinking (item C30- 40); (5) opportunity recognition skills (items O41-O54); and (6) teamwork skills (items T55-T66). A total of 66 items were built during the construction phase. The four-point Likert scale was used with the following options; 1-strongly disagree, 2-disagree, 4-agree and 5-strongly agree.

The data were analyzed based on the Rasch measurement model using Winstep software version 3.71. Testing was done based on item fit, unidimensionality, item polarity and local independence in explaining aspects of validity. While, aspects of item reliability were seen through the reliability index and separation index. The calibration scale which is one of the advantages of the Rasch model is also reported to test the suitability of the four-point Likert scale in METS.

Results and Discussions

Table 1 is showing the item fit of METS with MNSQ outfit values which ranges from 0.58 to 1.42. This values correspond to the item fit range of 0.5 to 1.5 logit (Boone et al., 2014). A total of 10 items were dropped because they fell out of MNSQ outfit range as shown in Table 2 below. Based from the findings, Zstd value was outside of -2.0 to +2.0 range. Zstd values are easily influenced by sample size. If the sample size is large, the value of Zstd is usually greater than 3.0. Therefore, experts like Sumintono dan Widhiarso (2015) suggested that Zstd value can be ignored if the MNSQ outfit and infit are good.

Table 1
Items Fit and Polarity Values of PTMEA

Item STATISTICS: MISFIT ORDER

```
-------------------------------------------------------------------------------------------
|ENTRY  TOTAL TOTAL        MODEL|  INFIT  |  OUTFIT  |PT-MEASURE |EXACT MATCH|    |
|NUMBER SCORE COUNT MEASURE S.E. |MNSQ ZSTD|MNSQ ZSTD|CORR. EXP.| OBS% EXP%| Item |
|---------------------------------+---------+---------+----------+-----------+------|
|  45  1379  455    .33   .09|1.43  5.2|1.42  5.1|A .44  .52| 61.7  67.2| T2  |
|  34  1100  455   2.18   .08|1.28  4.2|1.40  5.5|B .45  .57| 51.1  57.5| C2  |
|  21  1184  455   1.68   .08|1.28  4.0|1.40  5.3|C .44  .56| 53.1  59.0| R10 |
|   8  1493  455   -.61   .09|1.36  4.7|1.40  4.8|D .44  .50| 60.8  68.3| D8  |
|   3  1302  454    .88   .08|1.32  4.2|1.40  4.9|E .36  .54| 55.2  64.4| D3  |
|  30  1314  455    .81   .08|1.34  4.4|1.37  4.7|F .50  .53| 58.0  64.9| O9  |
|  44  1373  455    .37   .09|1.33  4.2|1.33  4.1|G .46  .53| 62.6  67.1| T1  |
|  48  1376  455    .35   .09|1.32  4.1|1.31  3.8|H .55  .52| 60.0  67.2| T5  |
|  17  1450  455   -.24   .09|1.26  3.3|1.31  3.8|I .47  .51| 64.6  68.6| R4  |
|  29  1375  455    .36   .09|1.26  3.3|1.27  3.4|J .49  .53| 61.1  67.1| O8  |
|  50  1366  455    .43   .09|1.26  3.3|1.22  2.8|K .53  .53| 64.4  66.9| T7  |
|  27  1411  455    .07   .09|1.24  3.1|1.23  3.0|L .53  .52| 62.2  68.1| O5  |
|  22  1421  455   -.01   .09|1.24  3.1|1.24  3.0|M .47  .52| 65.9  68.3| R13 |
|   6  1521  455   -.85   .09|1.21  3.0|1.24  3.0|N .46  .49| 60.4  68.0| D6  |
|  19  1434  455   -.11   .09|1.22  2.9|1.21  2.6|O .52  .52| 63.9  68.5| R6  |
|  32  1426  455   -.05   .09|1.19  2.5|1.18  2.3|P .55  .52| 62.8  68.3| O11 |
|  25  1451  455   -.25   .09|1.08  1.1|1.11  1.5|Q .56  .51| 67.9  68.6| O3  |
|  28  1301  455    .91   .08|1.06   .8|1.11  1.4|R .56  .54| 61.9  64.1| O6  |
|  38  1295  455    .95   .08|1.05   .7|1.09  1.2|S .45  .54| 62.6  63.8| C7  |
|  35  1271  455   1.11   .08|1.06   .9|1.08  1.1|T .49  .54| 64.2  62.9| C3  |
|  23  1487  455   -.56   .09|1.07  1.0|1.04   .5|U .57  .50| 67.7  68.5| O1  |
|  36  1219  455   1.46   .08|1.00   .1|1.06   .9|V .52  .55| 63.1  60.2| C4  |
|  14  1448  455   -.23   .09|1.06   .8|1.05   .6|W .53  .51| 64.4  68.6| P6  |
|  55  1551  455  -1.12   .09| .92 -1.2|1.05   .7|X .45  .48| 72.6  67.5| T12 |
|  51  1611  455  -1.68   .10|1.04   .7| .98  -.3|Y .42  .45| 72.1  67.3| T8  |
|   2  1377  455    .34   .09|1.01   .2|1.03   .5|Z .47  .52| 66.8  67.2| D2  |
|     BETTER FITTING OMITTED    +----------+----------+          |      |    |
|  13  1396  455    .19   .09|1.01   .2|1.00   .0|z .52  .52| 69.5  67.8| P5  |
|   4  1398  455    .18   .09|1.01   .1|1.00   .0|y .55  .52| 66.6  67.8| D4  |
|  24  1524  455   -.88   .09|1.00   .1| .99  -.1|x .54  .49| 69.9  67.9| O2  |
```

```
| 31 1360  455   .47    .09| .97 -.4| .98  -.3|w .53  .53| 67.3 66.7| O10 |
| 15 1487  455  -.56    .09| .94 -.9| .90 -1.4|v .54  .50| 71.2 68.5| R2  |
|  7 1493  455  -.61    .09| .93 -1.0| .93  -.9|u .51  .50| 71.5 68.3| D7  |
| 42 1493  455  -.61    .09| .91 -1.4| .88 -1.6|t .52  .50| 71.7 68.3| C11 |
| 26 1502  455  -.69    .09| .88 -1.7| .86 -2.0|s .62  .50| 73.0 68.2| O4  |
|  1 1452  455  -.26    .09| .88 -1.8| .88 -1.7|r .52  .51| 70.8 68.6| D1  |
| 52 1507  455  -.73    .09| .84 -2.4| .80 -2.8|q .58  .50| 75.2 68.2| T9  |
| 54 1505  455  -.71    .09| .80 -3.1| .77 -3.3|p .56  .50| 77.0 68.2| T11 |
| 16 1490  455  -.58    .09| .80 -3.0| .80 -2.9|o .56  .50| 72.8 68.4| R3  |
| 10 1414  455   .05    .09| .77 -3.4| .80 -2.9|n .53  .52| 75.7 68.2| P2  |
| 53 1538  455 -1.00    .09| .80 -3.4| .77 -3.3|m .55  .49| 76.5 67.7| T10 |
| 33 1329  455   .70    .09| .78 -3.4| .77 -3.4|l .54  .53| 72.6 65.5| C1  |
| 20 1409  455   .09    .09| .77 -3.4| .77 -3.3|k .58  .52| 76.5 68.1| R8  |
|  9 1379  455   .33    .09| .75 -3.7| .77 -3.4|j .51  .52| 73.7 67.2| P1  |
| 11 1410  455   .08    .09| .72 -4.3| .71 -4.4|i .57  .52| 75.7 68.1| P3  |
| 37 1424  455  -.03    .09| .71 -4.4| .70 -4.4|h .56  .52| 75.7 68.3| C5  |
| 43 1427  455  -.05    .09| .68 -4.9| .68 -4.8|g .57  .52| 77.4 68.4| C12 |
| 47 1460  455  -.33    .09| .65 -5.7| .65 -5.3|f .60  .51| 77.4 68.6| T4  |
| 39 1396  455   .19    .09| .64 -5.6| .64 -5.6|e .60  .52| 76.1 67.8| C8  |
| 12 1424  455  -.03    .09| .63 -5.8| .63 -5.7|d .58  .52| 77.9 68.3| P4  |
| 46 1522  455  -.86    .09| .62 -6.6| .60 -6.2|c .61  .49| 79.4 68.0| T3  |
| 49 1482  455  -.51    .09| .61 -6.5| .60 -6.3|b .63  .50| 82.1 68.5| T6  |
| 40 1350  455   .55    .09| .58 -6.8| .58 -6.8|a .61  .53| 80.1 66.3| C9  |
|-------------------------------+----------+----------+-----------+-----------+------|
| MEAN 1415.9 455.0   .00    .09| .99 -.3|1.00  -.2|       | 68.6 67.1|     |
| S.D.  93.4   .1   .72    .00| .23 3.3| .24 3.4|       | 7.1  2.4|     |
-------------------------------------------------------------------------------------
```

Table 2
Dropped Misfit Items

Item Code	Dimension	MNSQ Outfit	Zstd	PTMEA Corr
R7	Dare to Take Risk	3.10	9.9	0.09
R14	Dare to Take Risk	2.90	9.9	-0.37
C6	Creativity and Innovation	1.98	9.9	0.03
R9	Dare to Take Risk	1.83	9.9	0.30
R12	Dare to Take Risk	1.74	8.6	0.34
R15	Dare to Take Risk	1.77	9.2	0.38
C13	Creativity and Innovation	1.69	8.0	0.24
R1	Dare to Take Risk	1.60	7.2	0.45
R11	Dare to Take Risk	1.58	7.3	0.34
C14	Creativity and Innovation	1.54	6.8	0.42

The measurement of item polarity values are referred as PTMEA Corr in Table 1 ranged from 0.4 to 0.63 logit. The reported values are accurate to the values of polarity item range as stated by Boone et al. (2014) which was between 0.4 to 0.85. A total of 10 items were dropped because the PTMEA Corr values were below 0.4, namely items R7, R14, C6, R9, R12, R15, C13, R11 and C14. Item R14 had a negative value indicating that it contradicted the

variables (Linacre, 2005). Items from the construct of decision-making skills, problem-solving skills, opportunity identification skills, and teamwork skills were found to be in 0.5 to 1.5 range (Boone et al., 2014).

The Principal Component Analysis of Residual in Table 3 is showing the raw variance as explained by measure was 34.7% which is close to the prediction model of 34.8%. This finding indicates that unidimensionality value did not meet the suggested minimum value of 40% by Ariffin et al. (2013). This unidimensionality value proved that METS is inadequate in explaining what should be measured.

The unexplained variance in the first contrast was 5.2% that was lesser than 10%, in accordance with the recommendation of Linacre (2007). The findings show that the ratio of the raw variance explained by measure to the unexplained variance in the first contrast are 6.67:1, in line with the minimum value of 3:1 (Conrad et al. 2012; Menjivar 2016). The Eigen value obtained was 4.3 which is in accordance to the limit of Linacre (2005) that is lesser than 5.0 indicate there is no second dimension exist in measurement.

Table 3
Principal Component Analysis
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	84.2	100.0%	100.0%
Raw variance explained by measures	=	29.2	34.7%	34.8%
Raw variance explained by persons	=	13.5	16.1%	16.1%
Raw Variance explained by items	=	15.7	18.6%	18.7%
Raw unexplained variance (total)	=	55.0	65.3% 100.0%	65.2%
Unexplned variance in 1st contrast	=	4.3	5.2% 7.9%	
Unexplned variance in 2nd contrast	=	3.8	4.5% 6.9%	
Unexplned variance in 3rd contrast	=	2.8	3.3% 5.1%	
Unexplned variance in 4th contrast	=	2.5	3.0% 4.6%	
Unexplned variance in 5th contrast	=	2.0	2.4% 3.7%	

Table 4 is showing 10 matching items with residual correlation values ranging from 0.40 to 0.68. This range was found to be less than 0.7 indicating that these items did not strongly correlated. In other words, these items measured different aspects from one another. These correlation values comply with the local independence requirements as suggested by Linacre (2012).

Table 4
Correlation Values between Items

```
                    LARGEST STANDARDIZED RESIDUAL CORRELATIONS
                    USED TO IDENTIFY DEPENDENT Item
                    ---------------------------------
                    |CORREL-| ENTRY    | ENTRY    |
                    | ATION|NUMBER Ite |NUMBER Ite |
                    |-------+-----------+----------|
                    | .68 |   44 T1  |   45 T2  |
                    | .62 |   23 O1  |   24 O2  |
                    | .48 |   34 C2  |   35 C3  |
                    | .47 |   46 T3  |   47 T4  |
                    | .47 |   24 O2  |   26 O4  |
                    | .47 |   23 O1  |   26 O4  |
                    | .44 |   53 T10 |   54 T11 |
                    | .43 |   24 O2  |   25 O3  |
                    | .41 |   25 O3  |   26 O4  |
                    | .40 |   18 R5  |   19 R6  |
                    ---------------------------------
```

The calibration scale is showing that all of the six criteria mentioned by Linacre (2002) and Bond & Fox (2015) have been met. The first criterion is showing that the observation frequency for each category exceeded 10. Second, the frequency distribution was unimodal which is a form of probability curve with visible peak as shown in Figure 1.
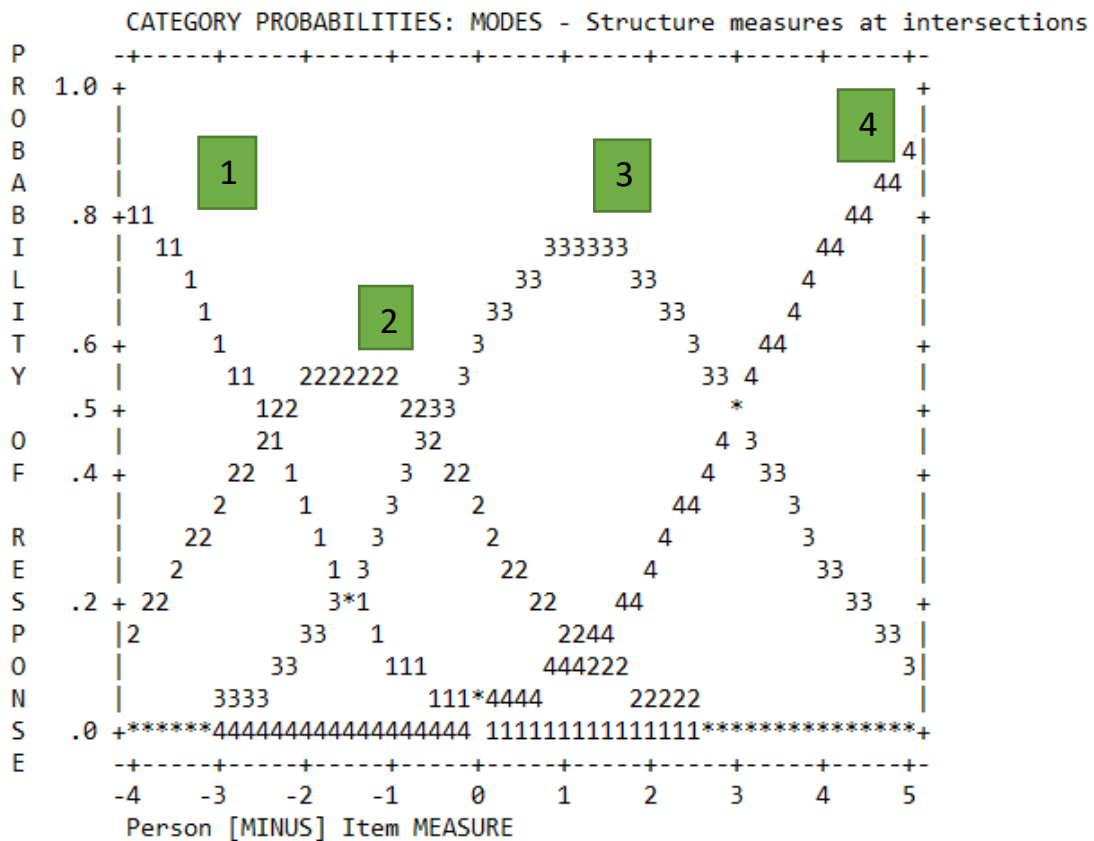

Figure 1. Four Categories Curve

Third, the value of the average measurement for the category needs to increase in line with the category scale. In this study, the four-point Likert scale (1<2<3<4) showed an increase in the measurement value of each scale which increased uniformly; -0.44 < 0.45 <1.49 < 2.94. Fourth, the MNSQ outfit value must be less than 2.0 which was also satisfied as presented in Table 5. The MNSQ outfit range for METS items was 0.91 to 1.48. The fifth criterion was also satisfied which includes the increase in threshold measurement value in line with the rating scale category. The threshold value increases starting from a scale of one to four which were -3.64 < -1.51 < 1.24 < 4.13. Sixth, the restrictions category must produce the value of one, but not exceeding five for four-point Likert scale, was also satisfied.

The Rasch analysis can help to validate the probability of uniformly distributed responses. Calibration structure is evaluated to confirm that the level classification is applicable. The separation value of 's' should be in the range of 1.4 < s < 5.0. If the value of s is less than 1.4, then the ratings will be combined. On the other hand, if the value of s is more than 5.0, then the ratings will be separated. The separation value, s for METS rating calibration scale lies in the range of 1.4 to 5.0 based on the calibration structure calculation done as shown in Table 5. The calculation for calibration structure is as follows.

$$S_{2-1} = (-2.44) - 0 = 2.44$$
$$S_{3-2} = (-.57) - (-2.44) = 1.87$$
$$S_{4-2} = 3.01 - (-.57) = 3.58$$

Table 5
Partial Credit Scale

Observation Category				Observation Average	Expected Sample	MNSQ Infit	MNSQ Outfit	Calibration Structure	Measurement Category
Label	Score	Observation Calculation	%						
1	1	315	1	-.44	-.95	1.31	1.48	-	-3.64
2	2	3165	13	.45	.47	1.00	1.00	-2.44	-1.51
3	3	14947	60	1.49	1.52	.92	.91	-.57	1.24
4	4	6597	26	2.94	2.89	.98	.97	3.01	4.13

All of the six criterias established by Linacre (2002) stated that the scale effectiveness testing has been satisfied, hence the four-point Likert scale is maintained. Typically, there are difficulties in meeting the six criteria of scale calibration, while sometimes some of the criteria cannot be met (Linacre, 2002). Although, the category absorption can help to improve its functionality, it does not mean that category scale conversion should be done (Andrich, 2002). This scenario explains that the scale effectiveness is done by observing the extent to which the criteria are met. Thus, scale calibration is preferred among most researchers to be used in testing especially when involving the developing of new instruments.

The individual reliability index is 0.94, while the reliability item index is 0.98. Reliability index values greater than 0.8 are considered as strong and acceptable (Bond & Fox 2015),

while 0.94 and above is considered as excellent. These findings indicate that the items are reliable, in addition to a sufficient number of samples (Linacre, 2012).

The individual separation index obtained was 3.93 while the separation item index was 7.64. Good separation item and individual exceeds the value of 2.0 (Linacre, 2018). In order to find the item strata and individual strata, the use of Strata = H is recommended. Strata values are calculated using formulas;

$H = \frac{(4G+1)}{3}$, when G is separation index (Wright & Masters, 2002)

Individual strata, $H = \frac{(4(3.93)+1)}{3} = 5.57$

Item strata, $H = \frac{(4(7.64)+1)}{3} = 10.52$

Strata calculations show that there are at least six levels of individual ability and 10 levels of item difficulty. Individual strata values and items with value more than five indicate an excellent separation (Fisher, 2007). Table 6 shows the summary of reliability, separation index and strata.

Table 6
Separation Index and Reliability

	Individu	Item	Cronbach Alpha (KR-20)
Reliability	0.94	0.98	0.95
Separation Index	3.93	7.64	-
Strata	5.57	10.52	-

In this study, a total of 66 items were being tested. From there, 10 items had to be dropped because they did not meet the psychometric characteristics of the Rasch model. Then, seven items from the 10 dropped items are considered for re-entry. The two items that were reintroduced were items R1 and C14 because the polarity values are in the range of 0.4 to 0.85 (Boone et al., 2014). Table 7 below is showing the items that were re-entered and those that were dropped.

Table 7
Items Dropped and Re-entered

Item Code	Action	Construct	MNSQ Outfit	PTMEA Corr
R7	Dropped	Dare to Take Risk	3.10	0.09
R14	Converted to positive item.	Dare to Take Risk	2.90	-0.37
C6	Dropped	Creativity & Innovation	1.98	0.03
R9	Dropped	Dare to Take Risk	1.83	0.30
R12	Converted to positive item.	Dare to Take Risk	1.74	0.34
R15	Converted to positive item.	Dare to Take Risk	1.77	0.38
C13	Converted to positive item.	Creativity & Innovation	1.69	0.24
R1	Maintained.	Dare to Take Risk	1.60	0.45
R11	Converted to positive item.	Dare to Take Risk	1.58	0.34
C14	Maintained.	Creativity & Innovation	1.54	0.42

The five items that were re-entered were negative items. The researcher re-entered the negative item after converting the negative item to positive. The justification of this initiative is that the mix of positive and negative items in a psychometric instrument will lower the validity and reliability of the instrument (Chyung et al. 2018). This is because the use of positive and negative words in one instrument could cause a confusion to the respondents when answering the question, especially when involving long instruments (DeVellis, 2017). The negative items that were converted to positive items are shown in Table 8. After considering the items to be re-included, the total number of items that will be used for the actual studynow became 63.

Table 8
Negative Items Converted to Positive Items

Item Code	Negative Statement	Positive Statement
R11	I avoid giving opinions so I would not be criticized.	I dare to express my opinion even when I might be criticized.
R12	I prefer to let other people make the decision for me.	I am responsible on the decision that I make.
C13	I do not like a routine task.	I like non-routine tasks.
R14	I am a perfectionist.	I am not a perfectionist.
R15	I always put off something that needs to be completed.	I will immediately complete any task eventhough it is difficult.

Conclusions

This study aims to test whether METS items have good psychometric characteristics among matriculation students based on Rasch analysis. Of the initial 66 items tested, 10 of which were dropped for failing to meet Rasch's assumptions and five items were reentered. Five of the items were found to be negative. The researcher suggested that further research should be carried out by reviewing the five items that were included after being converted to positive items. Besides that, further research should also be done to analyze the individual response by dropping out the misfit person. The purpose of this iniative is to test the effect of misfit person on unidimensionality. Since METS is a new instrument, there is a need to reexamine it from other aspects such as Differential Item Functioning (GDIF) in order to detect gender biasness. A specific testing is expected be able to produce METS as a robust psychometric instrument that can be utilized by all matriculation residence.

Acknowledgement

Corresponding Author

Amal Harun
Universiti Kebangsaan Malaysia. Selangor Malaysia
Email: p101201@siswa.ukm.edu.my

References

Ahmad, J., Moi, S. N., & Abdullah, S. (2020). Analisis Rasch bagi Ujian Pemikiran Sains Keusahawanan dalam Pendidikan STEM Sekolah Rendah. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *5*(10), 135-150.

Akbay, S. E., & Delibalta, A. (2020). Academic Risk Taking Behavior in University Students: Academic Procrastination, Academic Locus of Control, and Academic Perfectionism. *Eurasian Journal of Educational Research*, *89*, 159-177. DOI: 10.14689/ejer.2020.89.8

Aliu, J., & Aigbavboa, C. (2021). Reviewing Problem-Solving as a Key Employability Skill for Built Environment Graduates. In *Collaboration and Integration in Construction, Engineering, Management and Technology* (pp. 399-403). Springer, Cham.

Anwar, S. A. I. R. A., & Menekse, M. (2020). Unique contributions of individual reflections and teamwork on engineering students' academic performance and achievement goals. *International Journal of Engineering Education*, *36*(3), 1018-1033.

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*, *9*, 729-733.

Aziz, A. A., Masodi, M. S., & Zaharim, A. (2013). *Asas model pengukuran Rasch: pembentukan skala dan struktur pengukuran*. Penerbit Universiti Kebangsaan Malaysia.

Bacigalupo, M., Kampylis, P., Punie, Y., & Van den Brande, G. (2016). EntreComp: The Entrepreneurship Competence Framework. In Publications Office of the European Union. Publications Office of the European Union.

Blume, B. D., & Covin, J. G. (2011). Attributions to intuition in the venture founding process: Do entrepreneurs actually use intuition or just say that they do?. *Journal of Business Venturing*, *26*(1), 137-151. doi.org/10.1016/j.jbusvent.2009.04.002

Bond, T. G., & Fox, C. M. (2007). *Fundamental measurement in the human sciences*. Chicago, IL: Institute for Objective Measurement.

Bond, T. G., & Fox, C. M. (2015). Applying the Rasch model : fundamental measurement in the human sciences (3rd ed., Vol. 1). New York: Routledge. https://doi.org/10.1207/S15327574IJT013&4_10

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch Analysis in the Human Sciences. Springer. https://doi.org/10.1007/978-94-007-6857-4

Buang, N. A., Halim, L., & Meerah, T. S. M. (2009). Understanding the thinking of scientists entrepreneurs: Implications for science education in Malaysia. Journal of Turkish Science Education, 6(2), 3-11.

Conrad, K. M., Conrad, K. J., Dennis, M. L., & Funk, R. (2012). Validation of the Self Help Improvement Scale to the Rasch measurement model GAIN methods report 1.0. Chicago. Retrieved from http://gaincc.org/_data/files/Psychometrics_and_Publications/Working_Pa pers/Conrad_et_al_2011_SPS_Report.pdf

De Winnaar, K., and Scholtz, F. (2019), "Entrepreneurial decision-making: new conceptual perspectives", Management Decision, Vol. 58 No. 7, pp. 1283-1300. https://doi.org/10.1108/MD-11-2017-1152

DeVellis, R. (2017). *Scale development : theory and applications* (4th ed.). SAGE Publications Inc.

Dhliwayo, S. (2008). Experiential learning in entrepreneurship education. *Education+ training*. doi.org/10.1108/00400910810880560

Edwards-Schachter, M., García-Granero, A., Sánchez-Barrioluengo, M., Quesada-Pineda, H., & Amara, N. (2015). Disentangling competences: Interrelationships on creativity, innovation and entrepreneurship. *Thinking Skills and Creativity*. doi.org/10.1016/j.tsc.2014.11.006

Hancock, C., Hormiga, E., & Jaría-Chacón, N. (2020). Going it alone or Working as Part of a Team: The Impact of Human Capital on Entrepreneurial Decision Making.

Hnátek, M. (2015). Entrepreneurial thinking as a key factor of family business success. *Procedia- Social and Behavioral Sciences*, *181*, 342-348. doi.org/10.1016/j.sbspro.2015.04.896

Holm, H. J., Opper, S., & Nee, V. (2013). Entrepreneurs under uncertainty: An economic experiment in China. Management Science, 59(7), 1671-1687.

Ishak, Z., Buang, P., Aisyah, N., & Halim, P. (2013). *Pemikiran sains keusahawanan: Kesediaan mengintegrasi pemikiran keusahawanan dalam pengajaran sains di maktab rendah sains mara* (Doctoral dissertation, UKM).

Leffler, E. (2019). An Entrepreneurial Attitude: Implications for Teachers' Leadership Skills?. *Leadership and Policy in Schools*, 1-15. doi.org/10.1080/15700763.2019.1668021

Rajadurai, J., Sapuan, N. M., Daud, S., & Abidin, N. (2018). The marketability of technical graduates from higher educational institutions (HEIs) offering technical and vocational education and training (TVET): a case from Malaysia. *The Asia-Pacific Education Researcher*, *27*(2), 137-144.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. Journal of Applied Measurement, 3(1), 85–106.

Linacre, J. M. (2004). Test validity and Rasch measurement : construct, content etc. Rasch Measurement Transactions, 18(1), 970–971.

Linacre, J. M. (2005). Standard Errors : Means , Measures , Origins and Anchor Values. Rasch Measurement Transactions, 19(3), 1030.

Linacre, J. M. (2018). A user's guide to WINSTEPS: Rasch Model Computer Programs.

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. Academy of management Review, 21(1), 135-172.

Masters, G. N., & Wright, B. D. (2002). *The partial credit model. In Handbook of modern item response theory* (pp. 101-121). Springer, New York, NY.

Mohamad, A., Abdullah, S., Ishak, A. K., & Hashim, N. (2019). Entrepreneurship education as a way of cultivating entrepreneurial thinking among students of Malaysian public higher learning institutions. *International Journal of Social Science and Humanity*, *9*(4).

Othman N., & Ab Wahid, H. (2006). Ciri-ciri keusahawanan dan gaya pemikiran keusahawanan pengetua. Jurnal Pengurusan dan Kepimpinan Pendidikan, 16(01), 13-38.

Raosoft, I. (2016). Sample size calculator. 2004.

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan.* (B. Trim, Ed.) (1st ed.). Cimahi: Penerbit Trim 129 Komunikata.

Syukri, M., Halim, L., Meerah, T. S. M., & Buang, N. A. (2013). Pengetahuan Pedagogi Isi Kandungan Guru Sains Sekolah Rendah dalam Mengajarkan Pemikiran Sains Keusahawanan: Satu Kajian Kes. *Sains Humanika*, *63*(2). doi.org/10.11113/sh.v63n2.145

Timmons, J. A., Spinelli, S., & Tan, Y. (2004). *New venture creation: Entrepreneurship for the 21st century* (Vol. 6). New York: McGraw-Hill/Irwin.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. MESA press.