Test Instrument Validation in Plane Geometry using Rasch Model

Eyus Sudihartinih, Sufyani Prabawanto

Department of Mathematics Education, Universitas Pendidikan Indonesia
Email: eyuss84@upi.edu

Abstract
The purpose of this study was to describe the results of the analysis of the quality of the students' initial mathematical ability in the concept of plane geometry through the Rasch model. The benefit of this research is to provide knowledge about the validation of geometrical instruments with the Rasch model. This research is descriptive quantitative research. The study population was all students who attended two classes of analytic geometry at a university in Bandung. The sample was selected by purposive sampling so that one class of students was selected, consisting of 44 students studying analytic geometry (30 women and 14 boys). The research instrument was an essay that consisted of four geometry concepts. Based on the research results, it is known that all items meet the standard criteria as a measuring tool so that these questions can be used as instruments in further research..
Keywords: instrument validation, plane geometry, descriptive quantitative, Rasch model.

INTRODUCTION

Research on instrument testing plays an important role in data collection. The main indicators of the quality of research instruments are reliability and validity (Kimberlin & Winterstein, 2008). A valid instrument means that it can be used to measure what should be measured so that validity is an absolute requirement to produce valid research (Hidayati, 2012). A valid instrument is needed by a writer in research on geometry so that it needs the design and analysis of the instrument on the concept of geometry, especially the concept of plane geometry. It is because geometry has an important role to play. Volderman stated that geometry plays a role in our lives (Kambilombilo & Sakala, 2015); geometry plays a role in the concepts of astronomy, chemistry, biology, algebra, statistics, and calculus (Luneta, 2014).

In Indonesia, research on instrument validation has been carried out, including the Rasch model (Khumaeroh, Susongko, & M. Shaefur Rokhman, 2017; Nisa, Susongko, & Wikan Budi Utami, 2017; Purnomo, 2016; Susdelina, Perdana, & Febrian, 2018). The development of analysis using the Rasch model is one of the Theory Response Items that has been carried out since the 1960s by George Rasch (Rasch, 1968). This mathematical model was later popularized by Benjamin Wright and Geoff Masters (Wright & Masters, 1982). Data analysis with the Rasch model can be helped by Winstep software developed by Linacre (Linacre, 2006).

The advantage of the Rasch model is that it can determine the reliability and validity of research instruments (Bond & Fox, 2007; Razali & Shahbodin, 2016). The Rasch model can produce preferred and more precise measurement instruments (Sumintono & Widhiarso, 2014). The advantages of Rasch modeling can provide linear scales at the same interval, predict missing data, provide more precise estimates, detect

inaccurate models and produce replicable measurements (Sumintono, B.; and Widhiarso, W., 2014). Thus in this article explained about the validation of the instrument geometry problems using the Rasch model.

Research on the Rasch model for instrument validation has been carried out by several researchers (Khumaeroh et al., 2017; Maseko, Luneta, & Long, 2019; Nisa et al., 2017; Purnomo, 2016; Susdelina et al., 2018). Meanwhile, several other researchers studied the Rasch model for ability analysis (Folastri, Rangka, & Ifdil, 2017; Sari, Sekarwana, Hinduan, & Sumintono, 2016; Sudihartinih, Purniati, & Rohayati, 2019; Sudihartinih & Wahyudin, 2019a, 2019b; Widhiarso & Sumintono, 2016). In this study, instrument validation on plane geometry was not found.

RESEARCH METHOD

This research is descriptive quantitative research. The study population was all students who attended two classes of analytic geometry at a university in Bandung, Indonesia. The sample was selected by purposive sampling so that one class of students was selected, consisting of 44 students studying analytic geometry (30 girls and 14 boys). The instrument in this study was four problems in the concept of plane geometry. Three experts consulted questions before being tested on students. Here are the test instruments.

1. A segment A.B. with $A(3, -6)$ and $A(-5, -8)$ is known. Determine the distance of the midpoint of the segment to the line $3x - 4y = -8$.
2. Look for the point P which lies on the line through $P_1(2, -5)$ and $P_2(-3, 10)$ so $\left|\overrightarrow{P_1 P}\right| = 3\left|\overrightarrow{PP_2}\right|$.
3. Find the equation of a line through $(7, -3)$ intersect on coordinate axes at the intersection point of the same axis.
4. Find the equation of the circle whose center is at $3x - 5y = 8$ and offends its coordinate axis

The question is tested during the midterm student in two hours. Then the students' answers were given a score of 0-4. The scoring is as follows.

Table 1. Score

Information	Score
No answer	0
Can define graph	1
Can write the first equation	1
Can solve the first equation	1
Can write the second equation	1

Then the data were analyzed using the Rasch model with Winstep version 4.4.6. The steps are to analyze unidimensionality, person-item maps, item analysis, student's ability analysis, and instrument analysis.

1. Unidimensionality of measurement can be proven if Raw variance explained by measures $\geq 20\%$ (Note: general criteria for interpretation are: enough if $20 - 40\%$, good if $40 - 60\%$, and excellent if above 60%) and if Unexplained variance in 1st to 5th contrast of residuals $< 15\%$ each (Bambang Sumintono & Widhiarso, 2014).

2. The person-item map is above the average logit item, which is 0.00, which means that the average participant's ability is above the average item standard difficulty level.

3. For analysis items with criteria for checking the suitability of the item (item fit) or item mismatch (outliers or misfit) (Boone, Staver, & Yale, 2014) are as follows: (1) MNSQ OUTFIT value is greater than 0.5 and smaller than 1.5 and closer to 1 the better; (2) ZSTD OUTFIT value greater than -2.0 and smaller than $+2.0$ the closer to 0 the better; and (3) the value of PT MEASURE CORR is more than 0.4 and less than 0.85. An item can be considered fit if it meets at least 1 of the three criteria.

4. Analysis of students' abilities is by grouping them into high, medium, and high ability categories.

Instrument analysis, namely, analyzing the mean, standard deviation, separation, reliability, and Cronbach Alpha values.

RESULTS AND DISCUSSION

1. Unidimensionalitas
Unidimensionality analysis is needed to identify whether the instrument developed can measure what should be measured.

```
Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units
                                        Eigenvalue  Observed    Expected
Total raw variance in observations    =    9.3386 100.0%        100.0%
  Raw variance explained by measures  =    5.3386  57.2%         56.6%
    Raw variance explained by persons =    2.4552  26.3%         26.0%
    Raw Variance explained by items   =    2.8834  30.9%         30.6%
Raw unexplained variance (total)      =    4.0000  42.8% 100.0%  43.4%
  Unexplned variance in 1st contrast  =    1.5463  16.6%  38.7%
  Unexplned variance in 2nd contrast  =    1.3360  14.3%  33.4%
  Unexplned variance in 3rd contrast  =    1.1005  11.8%  27.5%
  Unexplned variance in 4th contrast  =     .0201    .2%   .5%
  Unexplned variance in 5th contrast  =     .0015    .0%   .0%

    STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
```

Figure 1. Unidimensionalitas

The results of data analysis in Figure 1 show that Raw variance explained by measures was observed at 57.2%, including the good category (Sumintono & Widhiarso, 2014). Whereas Unexplained variance in 1st to 5st contrast of residuals were 16.6%, 14.3%, 11.8%, 0.2%, and 0.0% respectively.

2. Analysis of Wright Map (Person-Item Map)
Wright map analysis can be seen in Figure 2.

```
MEASURE                      Person - MAP - Item
                                <more>|<rare>
  3                             15P  +
                                     |
                                     |
                                     |
                                   T |
                                     |
                                     |
                                     |
  2        12P  14P  19L  30L  37P  +
                                     |
                                     |
                     03L  11L  17L   |
                                   S |
    10P  18L  22P  24P  29L  44P   |T
                                     |
  1                      08P  41P  +
                                   |   N3
                                     |
                         07P  23P  |S
                                     |
         04L  05P  13L  26P  27P  M|
                                   |   N4
                                     |
    02P  31P  32L  34L  40P  42P   |
  0                                +M
                                   |   N2
                         16L  20P  36L  |
                                     |
    01P  09P  25P  28L  43P         |
                                     |
                                   S|S
             35P  38P  39P         |
                                     |
 -1                                +   N1
                                     |
                         21P        |
                                   |T
                                     |
                                     |
                                     |
                                   T|
                                     |
 -2                                +
                                     |
                         06P        |
                                     |
                                     |
                                     |
                                     |
                         33P        |
 -3                                +
                                <less>|<freq>
```
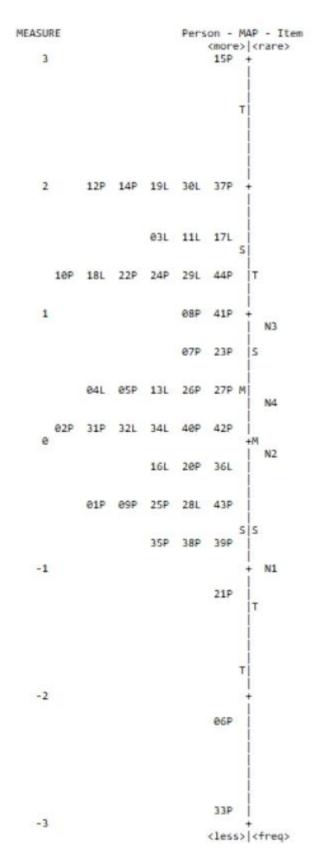
Figure 2. Variable map

Based on figure 2, it is known that the student's ability map spreads in a range of -1 to 2 logits. Students' abilities are at -3 logit to 3 logits. The position of N2 and N3 questions is between -SD and + SD so that the item items are within the student's ability while the N3 item items are above + SD so that they are above the student's ability. While the item N1 is under -SD so it has a difficulty level of items that are below the ability of students.

3. Item Analysis

This item analysis includes the level of difficulty (item measure), the level of suitability of item items (item fit), and the detection of item bias items.

2.1 Item Difficulty Level

The level of difficulty items can be examined in Figure 3: Item Measure.

```
Item STATISTICS:  MEASURE ORDER

---------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL          MODEL|  INFIT  |  OUTFIT  |PTMEASUR-AL|EXACT MATCH|       |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item |
|---------------------------------------+----------+----------+-----------+-----------+------|
|    3     73     44     .85      .18| .74 -1.36| .73 -1.26|  .69   .69| 46.5  41.0| N3   |
|    4     93     44     .26      .17|1.14  .71|1.11   .54|  .81   .71| 27.9  37.0| N4   |
|    2    106     44    -.13      .17| .97 -.06|1.11   .54|  .61   .72| 41.9  45.0| N2   |
|    1    133     44    -.99      .19|1.10  .50| .96 -.01|  .71   .69| 51.2  50.8| N1   |
|---------------------------------------+----------+----------+-----------+-----------+------|
| MEAN   101.3   44.0    .00      .18| .99  -.1| .98   .0|           | 41.9  43.4|       |
| P.SD    21.8    .0     .67      .01| .15   .8| .15   .7|           |  8.7   5.1|       |
---------------------------------------------------------------------------------
```

Figure 3. Item measure

From Figure 3, it is known that the SD value is 0.67. This SD value if combined with the average value of logit then the level of difficulty of items can be grouped into very difficult categories (greater +1 SD), hard categories (0.0 logit + 1 SD), easy categories (0.0 logit - 1 SD), and very easy categories (less than -1 SD). Thus, the limit value for the very difficult category is more than 0.67, the hard category is 0.00 to 0.67, the easy category is -0.67 to less than 0.00, and the very easy category is less than -0,67. Based on Figure 3 in sequence based on the level of difficulty (from the most difficult item to the easiest item), it is known that there is one item that is categorized as very difficult, namely item N3. The hard category is one item, N4. The easy category is one item, N2. While the category is very easy, there is 1 item, namely the question N1.

2.2 Item Match Level

Untuk melihat item yang berfungsi normal untuk pengukuran dapat ditinjau berdasarkan data pada Gambar 4: Item Fit Order yaitu kolom OUTFIT mean square (MNSQ), OUTFIT Z-standard (ZSTD), dan pengukuran titik (PT MEASURE CORR).
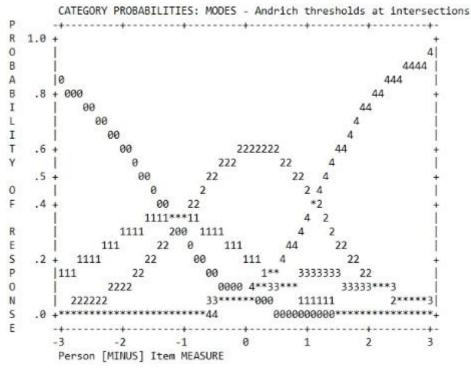
```
Item STATISTICS:  MISFIT ORDER

--------------------------------------------------------------------------
|ENTRY  TOTAL  TOTAL           MODEL|  INFIT  |  OUTFIT |PTMEASUR-AL|EXACT MATCH|        |
|NUMBER SCORE  COUNT  MEASURE   S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item  |
|-------------------------------------+----------+----------+-----------+-----------+-------|
|     4    93     44     .26    .17|1.14   .71|1.11   .54|A .81   .71| 27.9  37.0| N4    |
|     2   106     44    -.13    .17| .97  -.06|1.11   .54|B .61   .72| 41.9  45.0| N2    |
|     1   133     44    -.99    .19|1.10   .50| .96  -.01|b .71   .69| 51.2  50.8| N1    |
|     3    73     44     .85    .18| .74 -1.36| .73 -1.26|a .69   .69| 46.5  41.0| N3    |
|-------------------------------------+----------+----------+-----------+-----------+-------|
| MEAN  101.3   44.0     .00    .18| .99   -.1| .98   .0|           | 41.9  43.4|        |
| P.SD   21.8     .0     .67    .01| .15    .8| .15   .7|           |  8.7   5.1|        |
--------------------------------------------------------------------------
```

Figure 4. Misfit order

Based on criteria 1, 2, and 3, there are no misfit items (Boone, Staver, & Yale, 2014). Thus, all items of students' mathematical initial ability test items are declared fit in the sense of functioning normally and can be understood correctly by students and can measure what must be measured in this case is the initial mathematical ability.

2.3 Rating Scale Diagnostic
This diagnosis is carried out to determine whether participants have different answers in scores 0, 1, 2, 3, and 4.

Figure 5. Diagnostic

Differences in answers made by respondents if the observed average and Andrich Threshold values in Figure 5 show suitability and are equally increased in alternative answers 0, 1, 2, 3, 4. Thus it can be stated that students have answers on scores 0, 1, 2, 3, 4.

2.4 Detection of Bias Items

An item statement is said to contain bias if the probability value of the items, as listed in figure 6, is below 0.05 (Sumintono, B.; and Widhiarso, W., 2014). In the context of this study, bias can only be seen from the perspective of gender.

```
DIF class/group specification is: DIF=$S3W1

----------------------------------------------------------------------
| Person    SUMMARY DIF              BETWEEN-CLASS/GROUP Item         |
| CLASSES   CHI-SQUARED  D.F.  PROB.  UNWTD MNSQ  ZSTD  Number Name   |
|--------------------------------------------------------------------|
|    2        1.0751       1   .2998    1.1445    .57    1 N1         |
|    2        1.2554       1   .2625    1.3329    .68    2 N2         |
|    2        2.6685       1   .1024    2.9191   1.38    3 N3         |
|    2        3.2180       1   .0728    3.5469   1.59    4 N4         |
----------------------------------------------------------------------
```

Figure 6. Bias Item

The results of the analysis of bias based on gender note there is no single item that is biased. An overall picture of the logit position for each item by gender can be seen in the following figure.
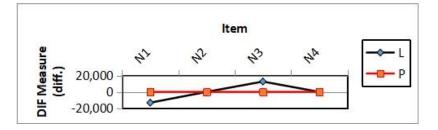

Figure 7. DIF measure

From the picture, it appears that each item can be worked out by male and female students.

4. Analysis of Student Ability

This analysis is carried out on two things, namely the level of individual ability (person measure) and the level of individual suitability (person measure).

4.1 Analysis of Individual Ability

Data on individual student's ability can be found in Person Measure From this figure. SD values are 1.19. This SD value when combined with an average logit (mean) value of 0.47 means that individual students' abilities can be grouped into the category of high ability (greater than 0.47 + 1.19 = 1.66), medium ability category (between 0.47 - 1.19 = -0.72 and 0.47 + 1.19 = 1.66 or -0.72 and 1.19), and the category of low ability (less than 0.47 - 1.19 = -0.72). Thus, the logit value limit for the high ability category is more than 1.66, the category of moderate ability from -0.72 to 1.66, and the category of low ability is less than -0.72.

```
|ENTRY   TOTAL  TOTAL          MODEL|  INFIT   |  OUTFIT  |PTMEASUR-AL|EXACT MATCH|       |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Person|
|---------------------------------------------------------------------------------------------|
|   15     16     4     3.42    1.53| MAXIMUM MEASURE  |  .00   .00|100.0 100.0| 15P  |
|   12     14     4     2.02     .66|1.85  1.17|1.97  1.02| -.11   .41| 25.0  49.5| 12P  |
|   14     14     4     2.02     .66|1.85  1.17|1.97  1.02| -.11   .41| 25.0  49.5| 14P  |
|   19     14     4     2.02     .66| .55  -.50| .38  -.19|  .74   .41| 50.0  49.5| 19L  |
|   30     14     4     2.02     .66| .55  -.50| .38  -.19|  .74   .41| 50.0  49.5| 30L  |
|   37     14     4     2.02     .66| .55  -.50| .38  -.19|  .74   .41| 50.0  49.5| 37P  |
|    3     13     4     1.64     .59|1.18   .47| .90   .22|  .74   .48| 25.0  31.5| 03L  |
|   11     13     4     1.64     .59| .21 -1.67| .20  -.96|  .89   .48| 50.0  31.5| 11L  |
|   17     13     4     1.64     .59|1.08   .33|1.03   .36|  .27   .48| 50.0  31.5| 17L  |
|   10     12     4     1.31     .56| .93   .10| .81   .00|  .74   .52| 50.0  42.1| 10P  |
|   18     12     4     1.31     .56| .92   .07| .79  -.03|  .54   .52| 50.0  42.1| 18L  |
|   22     12     4     1.31     .56|1.60  1.02|2.62  1.71| -.10   .52| 25.0  42.1| 22P  |
|   24     12     4     1.31     .56| .92   .07| .79  -.03|  .54   .52| 50.0  42.1| 24P  |
|   29     12     4     1.31     .56| .92   .07| .79  -.03|  .54   .52| 50.0  42.1| 29L  |
|   44     12     4     1.31     .56| .44  -.95| .39  -.77|  .83   .52| 50.0  42.1| 44P  |
|    8     11     4     1.01     .54| .36 -1.13| .33 -1.08|  .77   .55| 75.0  40.4| 08P  |
|   41     11     4     1.01     .54|1.07   .32| .92   .13|  .72   .55| 50.0  40.4| 41P  |
|    7     10     4      .72     .54| .54  -.59| .56  -.54|  .86   .55|  .0   35.3| 07P  |
|   23     10     4      .72     .54| .24 -1.47| .23 -1.42|  .85   .55| 50.0  35.3| 23P  |
|    4      9     4      .42     .54| .50  -.69| .51  -.62|  .62   .55| 50.0  45.8| 04L  |
|    5      9     4      .42     .54| .57  -.51| .58  -.47|  .77   .55| 50.0  45.8| 05P  |
|   13      9     4      .42     .54| .81  -.07| .80  -.07| -.22   .55| 50.0  45.8| 13L  |
|   26      9     4      .42     .54|2.70  1.95|2.66  1.87| -.52   .55| 50.0  45.8| 26P  |
|   27      9     4      .42     .54| .81  -.07| .80  -.07| -.22   .55| 50.0  45.8| 27P  |
|    2      8     4      .13     .55|1.44   .80|1.50   .85|  .66   .53| 25.0  41.2| 02P  |
|   31      8     4      .13     .55|2.44  1.76|2.54  1.78|  .20   .53|  .0   41.2| 31P  |
|   32      8     4      .13     .55| .92   .11| .92   .13|  .97   .53| 50.0  41.2| 32L  |
|   34      8     4      .13     .55| .92   .11| .92   .13|  .97   .53| 50.0  41.2| 34L  |
|   40      8     4      .13     .55| .58  -.51| .53  -.57|  .94   .53| 50.0  41.2| 40P  |
|   42      8     4      .13     .55| .31 -1.19| .35 -1.00|  .66   .53| 50.0  41.2| 42P  |
|   16      7     4     -.18     .56|1.39   .73|1.33   .66|  .85   .51| 50.0  39.6| 16L  |
|   20      7     4     -.18     .56| .45  -.81| .47  -.73|  .22   .51| 25.0  39.6| 20P  |
|   36      7     4     -.18     .56| .63  -.41| .66  -.33| -.11   .51| 25.0  39.6| 36L  |
|    1      6     4     -.50     .57| .11 -2.01| .12 -2.05|  .83   .50|100.0  44.9| 01P  |
|    9      6     4     -.50     .57| .52  -.61| .53  -.62|  .99   .50| 50.0  44.9| 09P  |
|   25      6     4     -.50     .57|1.27   .59|1.32   .65|  .46   .50|  .0   44.9| 25P  |
|   28      6     4     -.50     .57| .30 -1.21| .31 -1.21|  .54   .50| 50.0  44.9| 28L  |
|   43      6     4     -.50     .57|4.59  3.11|4.48  3.10| -.27   .50|  .0   44.9| 43P  |
|   35      5     4     -.84     .60|1.07   .33|1.20   .50|  .18   .49| 50.0  40.0| 35P  |
|   38      5     4     -.84     .60| .32 -1.15| .34 -1.12|  .89   .49| 50.0  40.0| 38P  |
|   39      5     4     -.84     .60|1.09   .36|1.07   .32|  .93   .49|  .0   40.0| 39P  |
|   21      4     4    -1.21     .63|1.59   .96|1.46   .81| -.46   .49| 25.0  47.2| 21P  |
|    6      2     4    -2.15     .77| .24 -1.24| .27  -.84|  .83   .42| 75.0  63.3| 06P  |
|   33      1     4    -2.90    1.01| .93   .29| .83   .33|  .11   .32| 50.0  71.9| 33P  |
|---------------------------------------------------------------------------------------------|
| MEAN    9.2    4.0     .47     .61| .98   .0| .98   .0|          | 41.9  43.4|       |
| P.SD    3.5     .0    1.19     .16| .80  1.0| .84  1.0|          | 20.7   7.1|       |
```

Figure 8. Person measure

In Figure 8, it is known sequentially based on ability level, it is known that six people are included in the high ability category, 32 people are in the medium ability category, and six people are in the low ability category.

4.2 Level of Individual Suitability

The suitability of individual responses based on their abilities can be examined based on the data in Figure 9, namely OUTFIT mean squire column (MNSQ), Z-standard OUTFIT (ZSTD), and point measuring correlation (PT MEASURE CORR).

```
|ENTRY   TOTAL  TOTAL          MODEL|   INFIT   |  OUTFIT  |PTMEASUR-AL|EXACT MATCH|        |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Person |
|-------------------------------------+----------+----------+-----------+-----------+--------|
|   43      6      4    -.50    .57|4.59  3.11|4.48  3.10|A-.27   .50|   .0  44.9| 43P    |
|   26      9      4     .42    .54|2.70  1.95|2.66  1.87|B-.52   .55| 50.0  45.8| 26P    |
|   22     12      4    1.31    .56|1.60  1.02|2.62  1.71|C-.10   .52| 25.0  42.1| 22P    |
|   31      8      4     .13    .55|2.44  1.76|2.54  1.78|D .20   .53|   .0  41.2| 31P    |
|   12     14      4    2.02    .66|1.85  1.17|1.97  1.02|E-.11   .41| 25.0  49.5| 12P    |
|   14     14      4    2.02    .66|1.85  1.17|1.97  1.02|F-.11   .41| 25.0  49.5| 14P    |
|   21      4      4   -1.21    .63|1.59   .96|1.46   .81|G-.46   .49| 25.0  47.2| 21P    |
|    2      8      4     .13    .55|1.44   .80|1.50   .85|H .66   .53| 25.0  41.2| 02P    |
|   16      7      4    -.18    .56|1.39   .73|1.33   .66|I .85   .51| 50.0  39.6| 16L    |
|   25      6      4    -.50    .57|1.27   .59|1.32   .65|J .46   .50|   .0  44.9| 25P    |
|   35      5      4    -.84    .60|1.07   .33|1.20   .50|K .18   .49| 50.0  40.0| 35P    |
|    3     13      4    1.64    .59|1.18   .47| .90   .22|L .74   .48| 25.0  31.5| 03L    |
|   39      5      4    -.84    .60|1.09   .36|1.07   .32|M .93   .49|   .0  40.0| 39P    |
|   17     13      4    1.64    .59|1.08   .33|1.03   .36|N .27   .48| 50.0  31.5| 17L    |
|   41     11      4    1.01    .54|1.07   .32| .92   .13|O .72   .55| 50.0  40.4| 41P    |
|   10     12      4    1.31    .56| .93   .10| .81   .00|P .74   .52| 50.0  42.1| 10P    |
|   33      1      4   -2.90   1.01| .93   .29| .83   .33|Q .11   .32| 50.0  71.9| 33P    |
|   18     12      4    1.31    .56| .92   .07| .79  -.03|R .54   .52| 50.0  42.1| 18L    |
|   24     12      4    1.31    .56| .92   .07| .79  -.03|S .54   .52| 50.0  42.1| 24P    |
|   29     12      4    1.31    .56| .92   .07| .79  -.03|T .54   .52| 50.0  42.1| 29L    |
|   32      8      4     .13    .55| .92   .11| .92   .13|U .97   .53| 50.0  41.2| 32L    |
|   34      8      4     .13    .55| .92   .11| .92   .13|V .97   .53| 50.0  41.2| 34L    |
|   13      9      4     .42    .54| .81  -.07| .80  -.07|u-.22   .55| 50.0  45.8| 13L    |
|   27      9      4     .42    .54| .81  -.07| .80  -.07|t-.22   .55| 50.0  45.8| 27P    |
|   36      7      4    -.18    .56| .63  -.41| .66  -.33|s-.11   .51| 25.0  39.6| 36L    |
|    5      9      4     .42    .54| .57  -.51| .58  -.47|r .77   .55| 50.0  45.8| 05P    |
|   40      8      4     .13    .55| .58  -.51| .53  -.57|q .94   .53| 50.0  41.2| 40P    |
|    7     10      4     .72    .54| .54  -.59| .56  -.54|p .86   .55|   .0  35.3| 07P    |
|   19     14      4    2.02    .66| .55  -.50| .38  -.19|o .74   .41| 50.0  49.5| 19L    |
|   30     14      4    2.02    .66| .55  -.50| .38  -.19|n .74   .41| 50.0  49.5| 30L    |
|   37     14      4    2.02    .66| .55  -.50| .38  -.19|m .74   .41| 50.0  49.5| 37P    |
|    9      6      4    -.50    .57| .52  -.61| .53  -.62|l .99   .50| 50.0  44.9| 09P    |
|    4      9      4     .42    .54| .50  -.69| .51  -.62|k .62   .55| 50.0  45.8| 04L    |
|   20      7      4    -.18    .56| .45  -.81| .47  -.73|j .22   .51| 25.0  39.6| 20P    |
|   44     12      4    1.31    .56| .44  -.95| .39  -.77|i .83   .52| 50.0  42.1| 44P    |
|    8     11      4    1.01    .54| .36 -1.13| .33 -1.08|h .77   .55| 75.0  40.4| 08P    |
|   42      8      4     .13    .55| .31 -1.19| .35 -1.00|g .66   .53| 50.0  41.2| 42P    |
|   38      5      4    -.84    .60| .32 -1.15| .34 -1.12|f .89   .49| 50.0  40.0| 38P    |
|   28      6      4    -.50    .57| .30 -1.21| .31 -1.21|e .54   .50| 50.0  44.9| 28L    |
|    6      2      4   -2.15    .77| .24 -1.24| .27  -.84|d .83   .42| 75.0  63.3| 06P    |
|   23     10      4     .72    .54| .24 -1.47| .23 -1.42|c .85   .55| 50.0  35.3| 23P    |
|   11     13      4    1.64    .59| .21 -1.67| .20  -.96|b .89   .48| 50.0  31.5| 11L    |
|    1      6      4    -.50    .57| .11 -2.01| .12 -2.05|a .83   .50|100.0  44.9| 01P    |
|-------------------------------------+----------+----------+-----------+-----------+--------|
| MEAN    9.2    4.0     .47    .61| .98   .0| .98   .0|           | 41.9  43.4|        |
| P.SD    3.5     .0    1.19    .16| .80  1.0| .84  1.0|           | 20.7   7.1|        |
```

Figure 9. Person Fit Order

Based on criteria (Boone, Staver, & Yale, 2014) is known that all students are declared fit in the sense of giving answers according to their level of ability.

5. Instrument Analysis

For instrument analysis, the information presented in figure: Summary Statistics are used.

```
SUMMARY OF 43 MEASURED (NON-EXTREME) Person
--------------------------------------------------------------------
|         TOTAL                        MODEL     INFIT       OUTFIT   |
|        SCORE    COUNT   MEASURE     S.E.    MNSQ  ZSTD   MNSQ  ZSTD |
|-------------------------------------------------------------------|
| MEAN     9.0      4.0      .41      .59     .98  -.04    .98   .01 |
|  SEM      .5       .0      .17      .01     .12   .16    .13   .15 |
| P.SD     3.4       .0     1.12      .08     .80  1.01    .84   .96 |
| S.SD     3.4       .0     1.13      .08     .81  1.02    .85   .98 |
| MAX.    14.0      4.0     2.02     1.01    4.59  3.11   4.48  3.10 |
| MIN.     1.0      4.0    -2.90      .54     .11 -2.01    .12 -2.05 |
|-------------------------------------------------------------------|
| REAL RMSE    .66 TRUE SD    .90 SEPARATION 1.36 Person RELIABILITY .65 |
|MODEL RMSE    .59 TRUE SD    .95 SEPARATION 1.59 Person RELIABILITY .72 |
| S.E. OF Person MEAN = .17                                          |
--------------------------------------------------------------------

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .67  SEM = 1.99

   SUMMARY OF 4 MEASURED (NON-EXTREME) Item
--------------------------------------------------------------------
|         TOTAL                        MODEL     INFIT       OUTFIT   |
|        SCORE    COUNT   MEASURE     S.E.    MNSQ  ZSTD   MNSQ  ZSTD |
|-------------------------------------------------------------------|
| MEAN   101.3     44.0      .00      .18     .99  -.05    .98  -.05 |
|  SEM    12.6       .0      .39      .00     .09   .47    .09   .42 |
| P.SD    21.8       .0      .67      .01     .15   .81    .15   .74 |
| S.SD    25.1       .0      .77      .01     .18   .93    .18   .85 |
| MAX.   133.0     44.0      .85      .19    1.14   .71   1.11   .54 |
| MIN.    73.0     44.0     -.99      .17     .74 -1.36    .73 -1.26 |
|-------------------------------------------------------------------|
| REAL RMSE    .18 TRUE SD    .64 SEPARATION 3.54 Item RELIABILITY .93 |
|MODEL RMSE    .18 TRUE SD    .65 SEPARATION 3.65 Item RELIABILITY .93 |
| S.E. OF Item MEAN = .39                                            |
--------------------------------------------------------------------
```

Figure 10. Summary Statistic

Based on Figure 10, the following information is known.

Table 2. Summary Statistic

	Mean	SD	Separation	Reliability	Cronbach Alpha
Person	0,41	1.12	1,36	0,65	0,67
Item	0,00	0.67	3.54	0,93	

Based on Table 2, a person measures 0.41 logit shows the average score of all participants working on items of initial student ability. The average value of a person that is greater than the average item (where the average item is 0.00 logit) shows that

the ability of participants is generally greater than the difficulty of the item items of the instrument. Cronbach Alpha value, which represents the interaction between person and item items as a whole, is 0.67, including enough category. Furthermore, the Person Reliability value is 0.65 as an indicator of the consistency of the respondents' answers, including the sufficient category. Item reliability of 0.93 as an indicator of the quality of the items on the instrument belongs to the very good category (Sumintono, B.; and Widhiarso, W., 2014).

Other data in a table that can be used are MNSQ INFIT and MNSQ OUTFIT, both in Table Person and Table Item. Based on the table, it is known that the average value of MNSQ INFIT and MNSQ OUTFIT are 0.98 and 0.98, respectively. Meanwhile, based on table item, it is known that the average value of INFIT MNSQ and OUTFIT MNSQ are 0.99 and 0.98, respectively. The criteria, the closer to number 1 the better, because the ideal value is 1 (Sumintono, B.; and Widhiarso, W., 2014). Thus, the average person and item approach the ideal criteria.

Meanwhile, related to INFIT ZSTD and OUTFIT ZSTD, the average values for the person, are -0.04 and 0.01, respectively. In contrast, the value of INFIT ZSTD and OUTFIT ZSTD for each item are -0.05 and -0.05. The ideal value of ZSTD is 0, the closer it is to 0, the better (Sumintono, B.; and Widhiarso, W., 2014). Thus it can be said that the quality of the person and items is good.

Thus it can be said that the quality of people and items is the latest when it comes to separation or grouping of people and items. Distinct separation shows how well a set of items in a student's geometry level thinking instrument spreads along with the range of logit skills. The greater the individual's separation, the better the instruments are arranged because the items in it can reach individuals with high to low levels of ability. In contrast, item separation shows how large the sample subject to measurement is spread along a linear interval scale. The higher the separation of items, the better the measurements are made. This index is also useful for defining the significance of the construct being measured.

In Figure 10, it is known that the separation for one person is 1.36, and one item is 3.54. The greater the separation value, the better the overall quality of the person and instrument. Separation values are calculated more precisely through the formula: $H = \{(4 \times separation) + 1\} / 3$. Thus the separation value for a person is 2.15 rounded to 2, while the separation for an item is 5.05 rounded to 5. This implies that the study participants have a variety of abilities that can be categorized into two groups. Meanwhile, the difficulty level items are spread out into five groups, from the easiest to the most difficult groups.

Based on information on measurement results, the picture is obtained, as shown in the following figure.
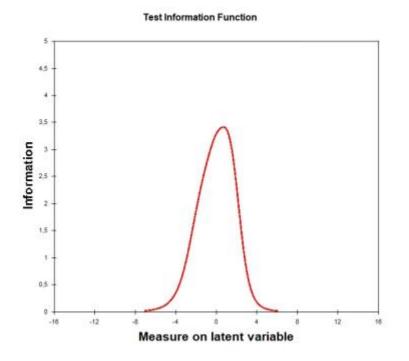
Figure 11. Measure on latent variable

The figure indicates that items about the level of geometrical thinking of students are more likely to produce great information on individuals with moderate levels of ability. Based on the findings, it is known that the instrument developed is valid so that this instrument will be able to measure the initial mathematical ability of the two-dimensional geometry concept. Although Rasch's analysis is very quantitative, it is clear that Rasch's analysis is also rich in qualitative (Boone, Townsend, & Staver, 2011). The first analysis shows the test instrument to have a good conceptual basis and be well targeted to groups, with a variety of items, so that students who have lower abilities can now answer a set of questions relatively easily. In contrast, while students with high ability skills will experience several things that are challenging (Maseko et al., 2019). The research can be continued with capability analysis (Folastri et al., 2017; Sari et al., 2016; Widhiarso & Sumintono, 2016).

CONCLUSION
All items meet the standard criteria as a measuring tool. The Cronbach Alpha value, which represents the interaction between the item person and the item as a whole is in the sufficient category. The Value of Person Reliability as an indicator of the consistency of respondents' answers is in the sufficient category. In contrast, item reliability as an indicator of the quality items on the instrument is classified as very good. Items are more likely to produce high levels of information about individuals of moderate ability. All students are declared fit in the sense of giving answers according to their level of ability, meaning that students are serious in giving answers. The position of the difficulty level is

one item that is categorized as very difficult, namely the item N3 question. The hard category is one item, N4. The easy category is one item, N2. While the very easy category has 1 item, namely question N1.

ACKNOWLEDGMENTS

REFERENCES

Bond, T. G., & Fox, C. M. (2007). *Applying The Rasch Model: Fundamental Measurement in the Human Sciences.* Mahwah, N.J.: Lawrence Erlbaum Associates Publishers.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Springer Dordrecht Heidelberg New York London.

Boone, W. J., Townsend, J. S., & Staver, J. (2011). Using Rasch Theory to Guide the Practice of Survey Development and Survey Data Analysis in Science Education and to Inform Science Reform Efforts: An Exemplar Utilizing STEBI Self-Efficacy Data. *Science Education*, *95*(2).

Folastri, S., Rangka, I. B., & Ifdil. (2017). Student's Self-concept Profile Based on Gender: a Rasch Analysis. *Advances in Social Science, Education and Humanities Research, Volume 118 9th International Conference for Science Educators and Teachers (ICSET)*.

Hidayati, K. (2012). Validasi Instrumen Non Tes dalam Penelitian Pendidikan Matematika. *Prosiding Jurusan Pendidikan Matematika*.

Kambilombilo, D., & Sakala, W. (2015). An Investigation into the Challenges In-Service Student Teachers Encounter in Transformational Geometry, "Reflection and Rotation". The Case of Mufulira College of Education. *Journal of Education and Practice*, *6*(2), 139–149.

Khumaeroh, S. U., Susongko, P., & M. Shaefur Rokhman. (2017). Penyusunan Skala Sikap Peserta Didik Terhadap Matematika Dengan Penerapan Model Rasch. *Jurnal Pendidikan MIPA Pancasakti*, *1*(1), 35–42.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. American *Journal of Health-System Pharmacy*, *65*(23), 2276–2284.

Linacre, J. M. (2006). A User's Guide to Winstep Ministep Rasch- Model Computer Programme. Retrieved from www.winstep.com.

Luneta, K. (2014). *Errors Displayed By Learners In The Learning Of Grade 11 Geometry*. *12*, 26–44.

Maseko, J., Luneta, K., & Long, C. (2019). Towards validation of a rational number instrument: An application of Rasch measurement theory. *Pythagoras*, *40*(1), a441.

Nisa, K., Susongko, P., & Wikan Budi Utami. (2017). Penyusunan Skala Minat Belajar Matematika Dengan Penerapan Model Rasch. *Jurnal Pendidikan MIPA Pancasakti*, *1*(1), 35–42.

Purnomo, S. (2016). *Pengembangan Soal Matematika Model PISA Konten Space dan Shape untuk Mengetahui Level Kemampuan Berpikir Tingkat Tinggi Berdasarkan Model Rasch*.

Rasch, G. (1968). A Mathematical Theory Of Objectivity And Its Consequences For Model Construction. *EUROPEAN Meeting on Statistics, Econometrics and Management Science, Amsterdam 2-7 September 1968*.

Razali, S. N., & Shahbodin, F. (2016). *Questionnaire on perception of online* collaborativelearning*: measuring validity and reliability using Rasch model.* Proceedings of the 4th International Conference on User Science and Engineering. Melaka, Malaysia, 1.

Sari, D. R., Sekarwana, N., Hinduan, Z. R., & Sumintono, B. (2016). Analisis Tingkat Kepuasan Masyarakat terhadap Dimensi Kualitas Pelayanan Tenaga Pelaksana Eliminasi Menggunakan Pemodelan Rasch. *JSK*, *2*(1).

Sudihartinih, E., Purniati, T., & Rohayati, A. (2019). Analisis Sikap Mahasiswa Calon Guru Matematika Dalam Perkuliahan Geometri Analitik Dengan Model Rasch. *KNPM 8 – IKIP SILIWANGI – Cimahi, 1 Agustus 2019*, 182–191.

Sudihartinih, E., & Wahyudin. (2019a). Analysis of students' self efficacy reviewed by geometric thinking levels and gender using rasch model. *Journal of Engineering Science and Technology*, *14*(1), 509–519.

Sudihartinih, E., & Wahyudin, W. (2019b). Pembelajaran Berbasis Digital: Studi Penggunaan Geogebra Berbantuan E-Learning Untuk Meningkatkan Hasil Belajar Matematika. *Jurnal Tatsqif*, *17*(1), 87–103. https://doi.org/10.20414/jtq.v17i1.944

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Cimahi, Indonesia: Trim Komunikata Publishing House.

Susdelina, Perdana, S. A., & Febrian. (2018). Analisis Kualitas Instrumen Pengukuran Pemahaman Konsep Persamaan Kuadrat Melalui Teori Tes Klasik Dan Rasch Model. *Jurnal Kiprah*, *6*(1), 41–48. https://doi.org/10.31629/kiprah.v6i1.574

Widhiarso, W., & Sumintono, B. (2016). Examining response aberrance as a cause of outliers in statistical analysis. *Personality and Individual Differences*, *98*, 11–15. https://doi.org/10.1016/j.paid.2016.03.099

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis Benjamin D. Wright Geofferey N. Masters*. Chicago: MESA PRESS.