Development of Critical Thinking Instrument with Minimum Competency Assessment Characters on Elasticity and Hooke's Law

Jayanti Eka Sari Ningsih[1], Dwi Susanti[1*], Vina Serevina [1], Slamet Maulana [2]

[1] Universitas Negeri Jakarta, Indonesia
[2]SMA Labschool Cibubur, Indonesia
*Corresponding author: jayantieka0@gmail.com

Abstract

Indonesia government sets a minimum competency assessment to prepare students in Indonesia for the 21st century and one of thinking skill that students must has been critical thinking skills. This study aims to develop an instrument test to measure critical skill that students have which are which followed by the minimum competency assessment for elasticity and Hooke's Law high secondary level. Furthermore, the instrument was developed on a research and development model and based on the development of an 8 steps test instrument by Mc Entire. The instrument test was validated by material experts who had been analyzed using Aiken V. This instrument tested on senior high school students and categorized as having a high reliability. Based on this, the measurement critical of thinking which is characterized by assessment competency minimum on elasticity and Hooke's law, in general can be done with this instrument. This research is beneficials as character-based essay tests to enhance critical thinking skills in high school students, focusing on elasticity and Hooke's law. It's intended as a tool to develop and refine these essential skills required for minimum competency assessments.

Key words: critical thinking, elasticity, Hooke's law, instrument test, minimum competency assessment

INTRODUCTION

Law Number 20 of 2003 from Indonesia Government about concerning the National Education System, Article 1 No.1 states that education is a conscious and planned effort to create a learning atmosphere and learning process so that students are active develop his potential to have spiritual religious power, self-control, personality, intelligence, noble character, and good skills needed by himself, society, nation and state. One of those skills what students need for themselves, society, nation and state is higher order thinking skills (Widyaningsih & Yusuf, 2018; Widyaningsih & Yusuf, 2018).

Higher Order Thinking Skills (HOTS) are the ability to connect, manipulate and transform existing knowledge and experience to think critically and creatively to solve problems in new situations. Higher Order Thinking Skills (HOTS) are abilities that are very necessary for daily life include the critical thinking ability. Critical thinking is the process needed to improve quality learners (Erfan & Ratu, 2018) and the most important indicators of student learning quality (Alsaleh, 2020). The ability to think critically can help to improve the affective, cognitive, and psychomotor aspects of students (As'ari, 2014; Ajwar, Prayitno, & Sunarno, 2015).

The development of science, technology and information in the 21st century requires capable human resources critical thinking (Kalelioğlu & Gülbahar, 2014). Critical thinking necessary to check the correctness of information, so we can decide the information deserves to be accepted or rejected (As'ari, 2014). The ability of student can be found out with process of learning and with the learning output that called by evaluate. In that process an instrument assessment is crucial. Assessment

can be used as a sign or evidence of an ongoing process so that it is comprehensive (Darmawan, Yatimah, Sasmita, & Syah, 2020). Assessment instruments such as test instruments must be developed with exact planning as a strategic step to solving physics problems (Gurcay & Gulbas, 2015).

Indonesia has changed the assessment from national exam to the assessment competency minimum in 2021 (Purnomo, Sa'dijah, Hidayanto, Permadi, & Anwar, 2022). The assessment of competency minimum that Indonesia use is needed for reach the target called by skills, one of which is critical thinking (Andiani, 2020). Minimum Competency Assessment is a fundamental competency assessment by all students to develop self-capacity and participate positively in society and the government makes this update in order to familiarize students with critical thinking that is contextual with their daily lives and avoid students from feeling tense in working on exam questions that contain only content in learning only (Ayuningtyas & Sukriyah, 2020). Several studies on Minimum Competency Assessment have been carried out, such as student readiness (Novrian, 2021), online learning (Chaichumpa & Temdee, 2018), pedagogy (Aunurrahman, 2020), critical thinking (Marunung, Hasanah, & Siswanto, 2020), design PISA, perceptions of prospective teachers (Novita, Mellyzar, & Herizal, 2021) and (Asfiati & Mahdi, 2020).

Assessment of minimum competency must carry out the student ability of the reflecting from the process learning. There are two competencies from namely literacy and the numeracy. The function of numeracy is for knowing the abilities from the student about concept of the unit educational (Maryuliana & Haviana, 2016; Marpadi, 2017). Numerical ability can be interpreted as a person's ability to formulate, apply, and interpret mathematics in various contexts, including the ability to do mathematical reasoning, and use concepts, procedures and facts to describe, explain or estimate phenomena/events (Ekowati, Astuti, Utami, Mukhlishina, & Suwandayani, 2019). Thus, the design of Minimum Competency Assessment Numeration questions is structured based on context in everyday life (Chaichumpa &

Temdee, 2018; Cahyanovianty & Wahidin, 2021) and world of work (Xiao, Barnard, Lan, & Burley, 2019). The assessment of numeracy skills is also related to science abilities. Science is a field of study related to physics.

Physics is a fundamental science (Oral & Erkilic, 2022), a branch of natural science which consists of concepts, principles, theories and or laws related to existing natural phenomena (Gurcay & Gulbas, 2015; Helmi, Rokhmat, & Ardhuha, 2017; Falani, 2020). One of the subjects in physics that is suitable for the education process in the 21st century is the elasticity material also Hooke's law since it is in accordance with the core competencies-3 in the 2013 curriculum which requires students to be able to understand, apply, analyze factual, conceptual, procedural knowledge and apply it to specific fields of study to solve the problem. The subject of elasticity and Hooke's law describes the physical phenomenon of elasticity. Elasticity and Hooke's law is actual material that is widely applied in everyday life, so this material is very important to be taught optimally. Elasticity and Hooke's law also have cognitive skills learning outcomes as well as many technological developments and their utilization.

In a preliminary study conducted via a Google form survey at SMAN 105 Jakarta, it was found that the written test by multiple choice is still an alternative at the assessment stage for both the Mid-Semester Examination and the Final Semester Examination with a percentage of 100% of 62 students results. This shows that the test by selecting answers is still an alternative in collecting data on the results of student competency achievement. In fact, the test by selecting this answer has weaknesses, namely it cannot measure students' thinking abilities in real terms, the possibility of students guessing the big answers and the occurrence of cheating (Istiyono, Mardapi, & Suparno, 2014).

Based on a survey that was conducted at SMAN 105 Jakarta, to avoid things that happened on the test by selecting answers, the selection was to carry out an assessment of students' abilities by using a supplier's test. According to (Marpadi, 2017) the description test is an instrument that can be used to carry out

measurements that aim to determine the competencies.

On the relevant research, it only measures the student ability to think critically test. There is still little research that measures the skill of thinking with critical with use the Indonesian government that called by assessment company minimum. The description of the minimum competency assessment character test in question is a description of the minimum competency assessment character test with items that measure students' critical thinking abilities in terms of numeracy. The description of the assessment competency minimum character test is designed to encourage the implementation of innovative learning that is oriented towards developing critical thinking skills and not focusing on memorization.

The purpose in this study is to develop an instrument test that can measure the student's critical thinking with the character of minimum competency assessment. This instrument is used for secondary level on elasticity and Hooke's law topics.

METHOD

The method that uses is the research and development model and based on the eight steps of a test instrument by Mc Entire as shown in Figure 1. The first step is defining, and the last step is compiling booklet it can be seen in Figure 1 (McIntire, 2000).
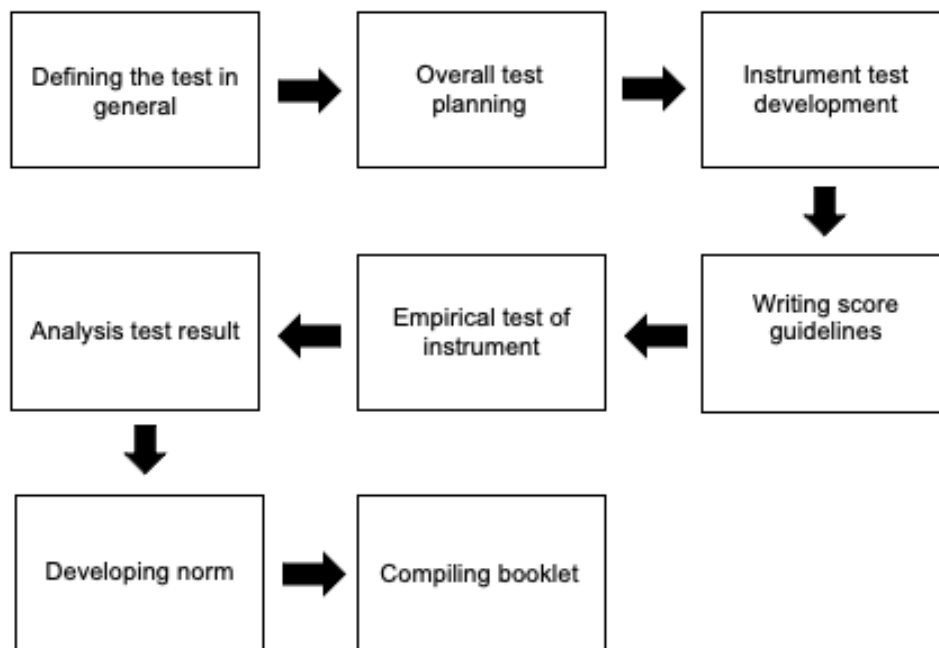
Figure 1. Research stages

The research subjects were students XI MIPA B at SMAN 105 Jakarta for the academic year 2022/2023. The instrument test developed with 10 items with seven aspects. A critical thinking instrument with the character of a minimum competency assessment for secondary students who have good validity and reliability certainly requires an appropriate testing method. The exact test is testing the validity and reliability through theoretical validity and empirical validity. The theoretical validation is carried out by validating theories or concepts in instruments carried out by expert practitioners, in this study the validation process involving three lecturers from Universitas Negeri Jakarta program Physics Education lecturers and two Physics teachers at SMAN 105 Jakarta. Moreover, for the empirical validation is about testing the students from XI MIPA B at SMAN 105 Jakarta. After the test, it will produce draft and the draff will be the final draft in this study.

Table 1. Indicators of critical thinking skill and the number of tests in each indicator (Facione, 1990)

Material	Cognitive Dimension	Aspek Critical Thinking Skills	Critical Thinking Skills Aspect Sub Indicators	Question Number
Elasticity	C5	Explanation	*Concluding Result*	1
	C5	Interpretation	*Interpreting Information*	2
	C5	Inference	*Constacting Premise*	3
	C5	Explanation	*Concluding Result*	4
Hooke's Law	C4	Inference	*Constacting Premise*	5
	C5	Evaluation	*Assessing Information Quality*	6
	C5	Explanation	*Presenting Arguments*	7
	C4	Analysis	*Considering Various Ideas*	8
	C5	Interpretation	*Interpreting Information*	9
	C5	Explanation	*Adjusting to Fact*	10

In this content validation using Aiken's validity coefficient, in this validity test was done quantitatively and qualitatively. The qualitative data provided is input and suggestions for the items being developed, then quantitative calculations are obtained with raw scores from experts totaling. In general, the assessment given is the suitability of basic competencies and indicators, suitability of indicators with material, use of language in test items, and construct tests (Falani, 2020). Aiken's V coefficient has a value range of 0 to 1 (Srirahayu & Arty, 2018). Can be calculated using the following equation:

$$V = \frac{\Sigma(ri-lo)}{n(c-1)} \quad (1)$$

The reliability analysis technique at this test phase uses a halving technique from the Cronbach-Alpha formula (Erika, Prahani, Supardi, & Tukiran, 2018), which will be analyzed using Winstep 5.1.2.0 according to the Rasch Model criteria, along with the reliability criteria:

Table 2. Reliability criteria

Score	Criteria
< 0.51	Lower reliability
0.50 – 0.61	Low reliability
0.60 - 0.71	Good reliability
0.70 - 0.81	High reliability
> 0.80	Higher reliability

Every measurement always contains errors. The source of measurement error comes from determining the exam material, the measuring party, the measuring party and the environment. In the process of analyzing these items, paying attention to the Rasch Model criteria using Winstep 5.1.2.0 The stages of item analysis carried out are as follows:

a. Grain discrimination

$$H = \frac{(4 \times \text{seperation}+1)}{3} \quad (2)$$

b. Difficulty Level

Table 3. Log Value Criteria

Logit Value	Explanation
>1	Very Difficult
0-1	Difficult
-1- 0	Easy
< -1	Very Easy

c. Item suitability level

Accepted Outfit Mean Square (MNSQ) value: 0.5< MNSQ<1.5 Accepted Outfit Z-Standard (ZSTD) value: -2 <ZSTD < 2. Point Measure Correlation (Pt Mean Corr) value: 0.4 < Pt Mean Corr < 0.85. In the suitability of this item, at least the item meets 2 conditions so that the item can be said to be fit.

d. Wright Map

It is an illustration of the distribution of students' abilities and the distribution of the difficulty levels of the items on the same scale. This map is divided into two parts, namely on the right the distribution of the difficulty levels of the items and on the left the distribution of student abilities.

RESULT AND DISCUSSION

All students had been taught the material a few days before the research was carried out as shown in Table 4.

Table 4. The respondents result of the reliabilities

```
SUMMARY OF 31 MEASURED Person
-------------------------------------------------------------------------------
|           TOTAL                        MODEL      INFIT       OUTFIT         |
|           SCORE    COUNT    MEASURE     S.E.    MNSQ   ZSTD   MNSQ   ZSTD     |
|-----------------------------------------------------------------------------|
| MEAN      38.1     10.0       2.53       .55     .99   -.07    .94   -.07     |
|  SEM       1.0      .0         .29       .02     .11    .22    .11    .18     |
| P.SD       5.4      .0        1.59       .11     .60   1.22    .62    .98     |
| S.SD       5.5      .0        1.61       .12     .61   1.24    .63   1.00     |
| MAX.      49.0     10.0       6.79      1.09    2.99   3.00   3.45   3.01     |
| MIN.      29.0     10.0        .06       .50     .14  -2.61    .11  -1.69     |
|-----------------------------------------------------------------------------|
| REAL RMSE    .62 TRUE SD    1.46 SEPARATION 2.35 Person RELIABILITY  .85     |
|MODEL RMSE    .57 TRUE SD    1.48 SEPARATION 2.61 Person RELIABILITY  .87     |
| S.E. OF Person MEAN = .29                                                    |
-------------------------------------------------------------------------------

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .86  SEM = 2.00
STANDARDIZED (50 ITEM) RELIABILITY = .97
```

In Table 4 the respondents average measure in the person is 2.53 (µ> 0.00), this indicates that in general the respondents have think with the critical skills with a good of characterized of assessment competency minimum. Furthermore, by calculating the separation person with the person strata equation.

It was found that with a separation index of 2.35 it would produce H = 3.46. Furthermore, the reliability value indicated by the Cronbach's alpha value which measures the interaction between respondents and items shows the result of r = 0.86 (Ekowati, Astuti, Utami, Mukhlishina, & Suwandayani, 2019; Wibisono, 2019; Falani, 2020).

tested again on students will produce almost the same results.

This instrument was based on the validation of material experts carried out by 3 physicists and 2 practicing experts, which were then calculated using Aiken's V to produce 10 valid question items with a range of Aiken's V index values of 0.86 - 0.90 where the standard V Aiken limit used is 0.80 with an error rate of 0.40. Reliability is the determination of test results where a test can be said to have a high level of confidence if the test is given to different subjects, it will produce fixed results (Wibisono, 2019). The table of coefficient V Aiken can be seen in Table 5.

Table 5. Coefficient V Aiken

No. of Items (m) or Raters (n)	Number of Rating Categories (c)											
	2		3		4		5		6		7	
	V	p	V	p	V	p	V	p	V	p	V	p
2							1,00	0,040	1,00	0,028	1,00	0,020
3							1,00	0,008	1,00	0,005	1,00	0,003
3			1,00	0,037	1,00	0,016	0,92	0,032	0,87	0,046	0,89	0,029
4					1,00	0,004	0,94	0,008	0,95	0,004	0,92	0,006
4			1,00	0,012	0,92	0,020	0,88	0,024	0,85	0,027	0,83	0,029
5			1,00	0,004	0,93	0,006	0,90	0,007	0,88	0,007	0,87	0,007
5	1,00	0,031	0,90	0,025	0,87	0,021	0,80	0,040	0,80	0,032	0,77	0,047

This explain that the data was very good fit, it can be clearly seen from the data between respondents and the instruments used and if

The reliability of the questions can be seen from the results of the analysis on the Winstep program using V Aiken's formula (Aiken,

1985). The results of the reliability of the test instrument can be seen in Table 6.

Table 6. Index of V Aiken

Question number	Index of V Aiken	Explanation
1	0.885	Valid
2	0.891	Valid
3	0.8675	Valid
4	0.903	Valid
5	0.877	Valid
6	0.891	Valid
7	0.891	Valid
8	0.897	Valid
9	0.9025	Valid
10	0.894	Valid

Based on the Table 7, the reliability of items got 0.95 with near to the model of the reliability of items with 0.96. That mean, the data was very good since the real score of reliability nearly to the model score of reliability.

From the Table 8, the most difficult items to work on is on the question number 3 with a measure value of 2.90 and the easiest of question to work is on the question number 5 with a measure value of -2.23 which is difference 4.67 point from the difficult to the easy questions. Furthermore, there are four similar score, which were numbers 6, 7, 9, and 10 with just difference of 0.10 of measure values.

Table 7. Item's reability

```
---------------------------------------------------------------------------
|          TOTAL                          MODEL      INFIT       OUTFIT    |
|          SCORE     COUNT     MEASURE     S.E.    MNSQ   ZSTD  MNSQ   ZSTD |
|-------------------------------------------------------------------------|
| MEAN     118.2     31.0         .00      .31    1.00   -.25   .94   -.25 |
|  SEM       5.7       .0         .51      .01     .17    .68   .16    .47 |
| P.SD      17.2       .0        1.53      .02     .50   2.05   .48   1.41 |
| S.SD      18.1       .0        1.61      .02     .53   2.16   .51   1.48 |
| MAX.     142.0     31.0        2.90      .35    2.00   3.21  1.91   2.16 |
| MIN.      86.0     31.0       -2.23      .29     .38  -3.18   .35  -2.35 |
|-------------------------------------------------------------------------|
| REAL RMSE    .34 TRUE SD   1.49  SEPARATION  4.41  Item   RELIABILITY .95|
|MODEL RMSE    .31 TRUE SD   1.49  SEPARATION  4.87  Item   RELIABILITY .96|
| S.E. OF Item MEAN =  .51                                                 |
```

Table 8. Item difficulty level table

```
---------------------------------------------------------------------------------------------
|ENTRY  TOTAL  TOTAL   JMLE   MODEL|  INFIT  | OUTFIT  |PTMEASUR-AL|EXACT MATCH|      |
|NUMBER SCORE  COUNT MEASURE  S.E. |MNSQ ZSTD|MNSQ ZSTD|CORR.  EXP.| OBS%  EXP%| Item |
|-------------------------------------------------------------------------------------------|
|   3     86    31    2.90    .32|1.13  .55|1.14  .58| .38  .69| 71.0  64.1| P3   |
|  10    102    31    1.38    .30| .62 -1.63| .58 -1.77| .74  .69| 83.9  59.3| P10  |
|   9    103    31    1.29    .29|1.18  .77|1.14  .58| .89  .69| 61.3  57.3| P9   |
|   2    105    31    1.12    .29| .43 -2.85| .47 -2.35| .87  .69| 80.6  58.4| P2   |
|   1    122    31    -.29    .29|1.19  .81|1.06  .29| .58  .68| 61.3  56.6| P1   |
|   8    125    31    -.54    .29|2.00  3.21|1.91  2.16| .50  .67| 45.2  58.2| P8   |
|   4    128    31    -.81    .30| .38 -3.18| .37 -2.08| .84  .66| 83.9  60.9| P4   |
|   7    134    31   -1.36    .31|1.32  1.20|1.46  1.00| .40  .62| 38.7  64.3| P7   |
|   6    135    31   -1.46    .32|1.33  1.23| .96  .08| .66  .62| 64.5  65.1| P6   |
|   5    142    31   -2.23    .35| .40 -2.61| .35 -1.00| .73  .54| 87.1  70.8| P5   |
|-------------------------------------------------------------------------------------------|
| MEAN  118.2  31.0    .00    .31|1.00  -.2| .94  -.3|           | 67.7  61.5|      |
```

Besides that, it can also be clearly seen that the reliability value of the items on the instrument, the results are in the Table 7 and 8.

The following is a table of suitability of items which can be seen from the Table 9.

Table 9. The suitability of the items

```
-------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL   JMLE   MODEL|  INFIT  |  OUTFIT  |PTMEASUR-AL|EXACT MATCH|       |
|NUMBER  SCORE  COUNT MEASURE  S.E. |MNSQ ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item  |
|-------------------------------------+---------+----------+-----------+-----------+------|
|     8    125    31    -.54    .29|2.00 3.21|1.91  2.16|A  .50  .67| 45.2  58.2| P8    |
|     7    134    31   -1.36    .31|1.32 1.20|1.46  1.00|B  .40  .62| 38.7  64.3| P7    |
|     6    135    31   -1.46    .32|1.33 1.23| .96   .08|C  .66  .62| 64.5  65.1| P6    |
|     1    122    31    -.29    .29|1.19  .81|1.06   .29|D  .58  .68| 61.3  56.6| P1    |
|     9    103    31    1.29    .29|1.18  .77|1.14   .58|E  .89  .69| 61.3  57.3| P9    |
|     3     86    31    2.90    .32|1.13  .55|1.14   .58|e  .38  .69| 71.0  64.1| P3    |
|    10    102    31    1.38    .30| .62 -1.63| .58 -1.77|d  .74  .69| 83.9  59.3| P10   |
|     2    105    31    1.12    .29| .43 -2.85| .47 -2.35|c  .87  .69| 80.6  58.4| P2    |
|     5    142    31   -2.23    .35| .40 -2.61| .35 -1.00|b  .73  .54| 87.1  70.8| P5    |
|     4    128    31    -.81    .30| .38 -3.18| .37 -2.08|a  .84  .66| 83.9  60.9| P4    |
|-------------------------------------+---------+----------+-----------+-----------+------|
| MEAN   118.2   31.0    .00    .31|1.00  -.2| .94   -.3|           | 67.7  61.5|       |
| P.SD    17.2    .0    1.53    .02| .50  2.0| .48   1.4|           | 15.9   4.2|       |
-------------------------------------------------------------------------------
```

From Table 9, for questions number 2, number 4 and number 8 tend to be unfit because the three questions do not meet the fit requirements according to the Rasch Model. The results of this analysis produce information on the distribution of items. Ability is an analysis that explains the abilities possessed by respondents who have taken part in trials of test instruments. The results of the analysis of the ability of the respondents can be seen in Table 10.

From Table 10, the results of the analysis related to the ability or ability to think critically with

Table 10. Critical thinking measured with assessment competency minimum

```
---------------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL   JMLE   MODEL|  INFIT  |  OUTFIT  |PTMEASUR-AL|EXACT MATCH|       |
|NUMBER  SCORE  COUNT MEASURE  S.E. |MNSQ ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Person|
|-------------------------------------+---------+----------+-----------+-----------+------|
|     2    49    10    6.79   1.09| .33  -.64| .11  -.64| .63  .39|100.0  89.2| 02P   |
|     3    48    10    5.91    .83| .97   .19| .35  -.16| .63  .51| 90.0  84.8| 03L   |
|     5    45    10    4.45    .61| .14 -2.61| .15 -1.52| .94  .67|100.0  72.2| 05P   |
|    16    44    10    4.10    .58|1.46   .98|1.52   .89| .59  .69| 80.0  70.3| 16P   |
|     1    43    10    3.77    .56| .43 -1.40| .61  -.56| .81  .71| 80.0  66.9| 01L   |
|     6    43    10    3.77    .56| .43 -1.40| .61  -.56| .81  .71| 80.0  66.9| 06P   |
|     8    42    10    3.46    .55|2.99  3.00|3.45  3.01| .60  .72| 60.0  61.4| 08P   |
|    25    42    10    3.46    .55|1.61  1.26|1.37   .79| .59  .72| 50.0  61.4| 25L   |
|    30    42    10    3.46    .55|2.20  2.09|1.77  1.34| .61  .72| 60.0  61.4| 30P   |
|     7    41    10    3.17    .53| .89  -.11| .78  -.30| .76  .72| 90.0  58.1| 07L   |
|     9    41    10    3.17    .53| .58  -.93| .58  -.84| .82  .72| 70.0  58.1| 09P   |
|    13    41    10    3.17    .53|2.01  1.88|1.67  1.29| .51  .72| 60.0  58.1| 13L   |
|    15    41    10    3.17    .53| .89  -.11| .78  -.30| .76  .72| 90.0  58.1| 15P   |
|    20    41    10    3.17    .53| .54 -1.08| .54  -.94| .80  .72| 70.0  58.1| 20P   |
|    22    39    10    2.63    .51| .69  -.65| .79  -.34| .69  .72| 60.0  59.5| 22P   |
|    23    39    10    2.63    .51| .43 -1.56| .41 -1.55| .86  .72| 80.0  59.5| 23P   |
|    27    39    10    2.63    .51| .43 -1.56| .41 -1.55| .86  .72| 80.0  59.5| 27L   |
|    14    37    10    2.12    .50|1.56  1.29|1.43  1.01| .49  .70| 50.0  57.1| 14L   |
|     4    36    10    1.88    .50|1.09   .36|1.04   .25| .56  .69| 50.0  56.4| 04P   |
|    17    36    10    1.88    .50|1.10   .37| .97   .08| .72  .69| 50.0  56.4| 17P   |
|    10    35    10    1.63    .50|1.19   .57|1.27   .72| .73  .68| 40.0  53.6| 10P   |
|    19    35    10    1.63    .50|1.28   .77|1.09   .35| .66  .68| 60.0  53.6| 19P   |
|    29    35    10    1.63    .50|1.38   .95|1.30   .76| .73  .68| 60.0  53.6| 29L   |
|    11    33    10    1.13    .50| .88  -.16| .89  -.11| .62  .66| 40.0  56.9| 11P   |
|    12    32    10     .88    .51| .96   .05| .90  -.09| .72  .66| 60.0  58.6| 12L   |
|    18    32    10     .88    .51| .96   .05| .90  -.09| .72  .66| 60.0  58.6| 18L   |
|    26    32    10     .88    .51| .62  -.88| .67  -.71| .62  .66| 60.0  58.6| 26P   |
|    31    31    10     .61    .52| .87  -.17| .98   .10| .59  .65| 50.0  58.1| 31P   |
|    24    30    10     .34    .53| .37 -1.72| .37 -1.69| .81  .64| 80.0  60.0| 24L   |
|    21    29    10     .06    .54| .70  -.55| .76  -.39| .69  .64| 70.0  60.7| 21L   |
|    28    29    10     .06    .54| .70  -.55| .76  -.39| .69  .64| 70.0  60.7| 28P   |
|-------------------------------------+---------+----------+-----------+-----------+------|
| MEAN   38.1   10.0   2.53    .55| .99   -.1| .94   -.1|           | 67.7  61.5|       |
| P.SD    5.4    .0    1.59    .11| .60  1.2| .62   1.0|           | 16.2   7.9|       |
---------------------------------------------------------------------------------------
```

distribution of respondents' abilities and the the the assessment competency minimum. In

determining students' abilities, they are not reviewed based on the raw scores obtained but also look at the level of suitability of the analysis results, namely the MNSQ, ZSTD, Pt.Corr scores for each individual. For the highest ability score was obtained by 02 P with an ability value of 6.79 while the students who obtained the lowest ability score were students 21 L and 28 P with the same ability value of 0.06.Students with numbers 09 P, 13 L, 15 P, and 20 P had scores that appears the most with 42 total score.

Table 11 show the results of the analysis of the respondents' abilities and the consistency.

Table 11. Scalogram

```
GUTTMAN SCALOGRAM OF RESPONSES:
Person |Item
       |          1
       |5674812903
       |----------
    2 +5555555554   02P
    3 +5555555553   03L
    5 +5555554443   05P
   16 +5555354543   16P
    1 +5545544443   01L
    6 +5545544443   06P
    8 +5535554541   08P
   25 +5554354533   25L
   30 +5555354523   30P
    7 +5555344433   07L
    9 +5545543433   09P
   13 +5545524443   13L
   15 +5555344433   15P
   20 +5544544433   20P
   22 +4544543343   22P
   23 +5544543333   23P
   27 +5544543333   27L
   14 +4544523343   14L
    4 +5354334333   04P
   17 +4354543332   17P
   10 +4554342233   10P
   19 +4353543332   19P
   29 +4354543232   29L
   11 +5343343233   11P
   12 +4533343232   12L
   18 +4533343232   18L
   26 +4344333233   26P
   31 +4343342233   31P
   24 +4343333232   24L
   21 +4343332223   21L
   28 +4343332223   28P
```

From the Table 11 it can be clearly seen the concluded of the respondents who has the highest ability were the student number 02 P

because she has consistency when answering the questions with given compared to respondent number 03 L. The consistence can be seen on the score of question number 3, the score of student of 02 P was 4. On the other hand, the score of student number 03 L was 3. On the flip hand, the student number 02 P got highest score with 49 points and the student number 03 L got 48 points.

From the reliability of the test instrument to measure critical thinking ability, it obtained a fairly good score, namely the Alpha Cronbach (KR-20) value of 0.86, so that it can be said that the interaction between the respondent and the test instrument is very reliable. Furthermore, this is also supported by the ability of the instrument which has a separation value of H = 3.46. This indicates that this instrument can classify the ability of respondents in two groups of high and low ability, then on the reliability item a value of 0.95 is obtained which indicates that the instrument has a very good reliability value and with a separation value of H = 6.21 which explains that this item instrument can divide the questions into 5 levels of questions, namely very easy questions, easy questions, moderately difficult questions, difficult questions, very difficult questions. Furthermore, if we look at the item suitability analysis for questions number 2, number 4 and number 8, misfit according to the Rasch model is not due to meeting at least two of the criteria that must be filled in, so these three questions must be eliminated.

The division of the level of questions can be shown by the value of the level of difficulty in which the value of the level of difficulty is obtained from -2.23 to 2.90. for the most difficult question is question number 3 which is proven that this question was only able to be done by 14 respondents, this was most likely due to the complicated explanation of the questions so that the respondents could not understand and confirm the answers that had to be written down. then for the easiest question, namely question no 5, it shows that all students are able to answer the question. This is probably the form of questions that are too common place and are often found on tests or in school text books. Wright Map state that all students in general can solve the questions given and from the results of this description it can be said that the questions in

general are on average below the ability of the test students. The test instrument is achieved optimally when the average item is aligned with

developed instrument is still very good because it meets the requirements of the Rasch model as displayed in Figure 2.
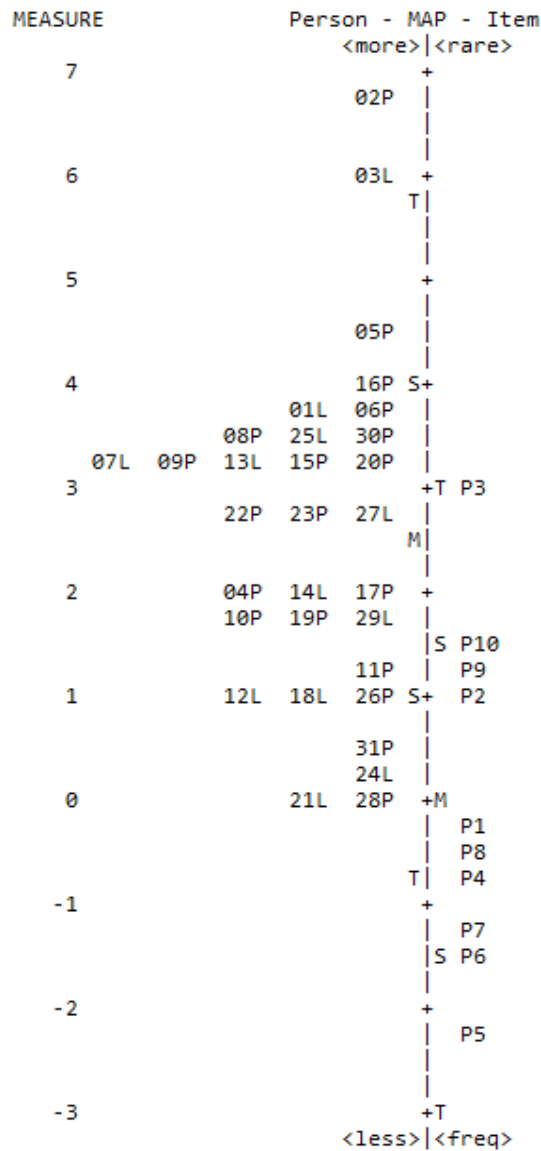
```
    MEASURE                    Person - MAP - Item
                                 <more>|<rare>
       7                             +
                                02P  |
                                     |
                                     |
       6                        03L  +
                                   T |
                                     |
                                     |
       5                             +
                                     |
                                05P  |
                                     |
       4                        16P S+
                        01L  06P     |
                   08P  25L  30P      |
         07L  09P  13L  15P  20P     |
       3                            +T P3
                   22P  23P  27L     |
                                   M |
                                     |
       2             04P  14L  17P   +
                     10P  19P  29L    |
                                    |S P10
                          11P  |     P9
       1             12L  18L  26P S+ P2
                                     |
                               31P  |
                               24L  |
       0             21L  28P  +M
                                  |  P1
                                  |  P8
                               T |  P4
      -1                            +
                                  |  P7
                                 |S  P6
                                  |
      -2                            +
                                  |  P5
                                  |
                                  |
      -3                          +T
                                 <less>|<freq>
```

Figure 2. Wright Map

the average respondent's ability. However, the

CONCLUSION

The results of the data analysis of the developed instruments can be used as supp instrument with the characteristics assessment competency minimum for the elasticity material ort to state that the critical thinking measurement and hook law in secondary level is feasible to use. For the result of this study, it was concluded that the instrument met the unidimensional

requirement that called by constructive measure, in this case, the critical thinking skills abilities with character assessment of competency minimum. Moreover, the items gave result that the instrument has discriminating power and with that standard, the instrument test can be develop in this article and the instrument can be used to measure the critical skill of the students with characteristic of the assessment competency

minimum that Indonesia government has for the elasticity and hook law material.

ACKNOWLEDGMENTS

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*, 131-142.

Ajwar, M., Prayitno, B., & Sunarno, W. (2015). Pengaruh Pembelajaran Inkuiri Terbimbing Dan Inkuiri Bebas Termodifikasi Terhadap Prestasi Belajar Ditinjau Dari Berpikir Kritis Dan Kedisiplinan Belajar Siswa Kelas X MIA SMA Negeri 8 Surakarta Tahun Pelajaran 2014/2015. *Jurnal Inkuiri, 4(3)*, 127-135.

Alsaleh, N. J. (2020). Teaching Critical Thinking Skills. *Turkish Online Journal of Educational, 19(1)*, 21-39.

Andiani, D. (2020). Analysis of the Design of the Minimum Competency Assessment (AKM) Numeration of the Independent Learning Program. *Journal of Mathematics and Mathematics Education (Majamath). 4(1)*, 25-35.

As'ari, A. R. (2014). Ideas for Developing Critical Thinking at Primary School Level. *Paper presented at an International Seminar on Addresing Higher Order Thinking Skills: Critical Thinking Issues In Primary Education.* Makasar: Islamic University Of Muhammadiyah Makassar.

Asfiati, & Mahdi, N. (2020). on Children with Special Needs. *J. Islamic Early Child. Educ. 3(2)*, 59.

Aunurrahman. (2020). Pedagogy. *J. Educ. Teach. Learn. 5(1)*, 296.

Ayuningtyas, N., & Sukriyah, D. (2020). Analisis Pengetahuan Numerasi Mahasiswa Matematika Calon Guru. *Matematika Dan Pendidikan Matematika, 9(2)*, 237–247.

Cahyanovianty, A. D., & Wahidin, W. (2021). Analisis Kemampan Numerasi Peserta Didik Kelas VIII dalam Menyelesaikan Soal Asesmen Kompetensi Minimum (AKM). *Jurnal Pendidikan Matematika, 5(2)*, 1439-1448.

Chaichumpa, S., & Temdee, P. (2018). Online Learning. *Int. J. Innov. Learn 23(1)*, 19-22.

Darmawan, D., Yatimah, D., Sasmita, K., & Syah, R. (2020). Analysis of Non-Formal Education Tutor Capabilities in Exploring Assessment for Science Learning. *Jurnal Pendidikan IPA Indonesia, 9(2)*, 267–275.

Ekowati, D. W., Astuti, Y. P., Utami, I. W., Mukhlishina, I., & Suwandayani, B. (2019). Literasi Numerasi di SD Muhammadiyah. *Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar, 3(1)*, 93.

Erfan, M., & Ratu, T. (2018). Pencapaian HOTS (Higher Order Thinking Skills) Mahasiswa Program Studi Pendidikan Fisika FKIP Universitas Samawa. *Jurnal Pendidikan Fisika dan Teknologi, 4(2)*, 208-212.

Erika, F., Prahani, B. K., Supardi, Z. A., & Tukiran. (2018). Development of graphic organizer-based argumentation learning (GOAL) model for improving the ability to argue and self-efficacy of chemistry teacher candidates. *World Trans on Engineering and Technology Education, 16(2)*, 179-185.

Facione, P. A. (1990). *Critical Thinking A Statement of expert consensus for purposes of educational assessment and instruction.* Millbrae: The California Academic Press.

Falani, I. (2020). The Combination Model of Respondents Theory Scored Items in the Mixed Format Test. *State University of Jakarta, 8(1)*, 7-14.

Gurcay, D., & Gulbas, E. (2015). Development of Three-Tier Heat, Temperature and Internal Energy Diagnostic Test. *Research in Science and Technological Education, 33(2)*, 197–217.

Helmi, F., Rokhmat, J., & Ardhuha, J. (2017). Pengaruh Pendekatan Berpikir Kausalistik dengan Modified Scaffolding Type 2B Berbantuan Lembar Kerja Terhadap Kemampuan Pemecahan

Masalah Fluid Dynamic Siswa. *Jurnal Pendidikan Fisika dan Teknologi. 3(1)*, 68-75.

Istiyono, E., Mardapi, D., & Suparno, S. (2014). Development of Physics Higher Order Thinking Ability Test (Pyshots) for High School Students. *Journal of Educational Research and Evaluation, 18(1)*, 1–12.

Kalelioğlu, F., & Gülbahar, Y. (2014). The Effect of Instructional Techniques on Critical Thinking and Critical Thinking Dispositions in Online Discussion. *Journal Educational Technology and Society, 17(1)*, 248-258.

Marpadi, D. (2017). *Measurement, Assessment and Evaluation of Education (Revised Edition).* Indonesia: Parama Publishing.

Marunung, A. S., Hasanah, M., & Siswanto, W. (2020). Critical Thinking. *J. Pendidik. Teor. Penelit. Pengemb. 5*, 730.

Maryuliana, S. I., & Haviana, S. F. (2016). Questionnaire Information System for Measuring the Scale of Needs for Additional Learning Materials to Support Decision Making in High Schools Using a Likert Scale. *Journal of Basic Educatio, 3(2)*, 27-44.

McIntire, S. A. (2000). *Foundation of psychological testing.* Boston: McGraw-Hill.: McGraw-Hill.

Novita, N., Mellyzar, M., & Herizal, H. (2021). design PISA, perceptions of prospective teachers. *J. Ilmu Sos.Pendidik. 5*, 172.

Novrian, S. P. (2021). Several studies on Minimum Competency Assessment have been carried out. *Jurnal Pendidikan, Sejarah dan Ilmu-ilmu Sosial 5(1)*, 15-20.

Oral, I., & Erkilic, M. (2022). Investigating the 21st Century Skills of Undergraduate Students: Physics Success, Attitude, and Perception. *Journal of Turkish Science Education, 19(1)*, 284-301.

Purnomo, H., Sa'dijah, C., Hidayanto, E., Permadi, H., & Anwar, L. (2022). Development of Instrument Numeracy Skills Test of Minimum Competency Assessment (MCA) in Indonesia. *International Journal of Instruction, 15(3)*, 635-648.

Srirahayu, R. R., & Arty, I. S. (2018). Validity and Reliability of Stem-Based Science Literacy Performance Assessment Instruments. *Journal of Educational Research and Evaluation, 22(2)*, 168–181.

Wibisono, S. (2019). Rasch Model Application for Validation of Religious Fundamentalism Measurement Instruments for Muslim Respondents. *Indonesian Journal of Psychology and Education Measurement, 3(3)*, 62-65.

Widyaningsih, S. W., & Yusuf, I. (2018). Project Based Learning Model based on Simple Teaching Tools and Critical Thinking Skills. *Physics Education Journal, 1(1)*, 12-21.

Xiao, F., Barnard, B. L., Lan, W., & Burley, H. (2019). Examining problem-solving skills in technology-rich environments as related to numeracy and literacy. *International Journal of Lifelong Education, 38(3)*, 327–338.