 doi: 10.21831/cp.v38i2.23062

A TEST OF ANALYTICAL THINKING AND CHEMICAL REPRESENTATION ABILITY ON 'RATE OF REACTION' TOPIC

Antuni Wiyarsi[1*], Atina Rizanatul Fachriyah[2], Didi Supriadi[3], Muhd Ibrahim bin Muhamad Damanhuri[4]
[1]Universitas Negeri Yogyakarta, [2]Sunan Kalijaga Islamic State University, Indonesia
[3]Universitas SarjanawiyataTamansiswa, Indonesia
[4]Universitas Pendidikan Sultan Idris, Malaysia
*e-mail: antuni_w@uny.ac.id

Abstract: Assessments play an important role in chemistry learning and for specific uses. The construction of a test based on multiple representation approaches is needed for measuring the 21st century thinking skills. This study aims to construct and validate a standardized test to measure students' analytical thinking and chemical representation ability in rate of reaction topic. The test captures four aspects on analytical thinking and four levels of multiple representations (macroscopic, sub-microscopic, symbolic and mathematic). A group of experts confirmed the construct and face validity of the Test of Analytical Thinking based on Multiple Representation (TAT-MR) with 32 items. The TAT-MR was then validated by participating 449 high school students. The test characteristics were analyzed usingRasch model with Partial Credit Model-1 Parameter Logistic (PCM-1PL) approach. The results of theRaschmodeling show that there are 22 TAT-MR items with excellent reliability. Hence, the TAT-MR is acceptable as a good instrument to collect the data. This study suggests that TAT-MR will prove to be a useful instrument for measuring the students' ability on analytical thinking and chemical representation for rate of reaction topic in chemistry learning.

Keywords: ***analytical thinking, chemical representation, Rasch model, rate of reaction, validation***

TEST BERPIKIR ANALISIS DAN KEMAMPUAN REPRESENTASI KIMIA PADA TOPIK LAJU REAKSI

Abstrak: Penilaian memegang peranan penting dalam pembelajaran kimia dan untuk penggunaan khusus. Penyusunan tes berbasis pendekatan multipel representasi diperlukan untuk mengukur keterampilan berpikir abad XXI. Penelitian ini bertujuan untuk membuat separangkat soal yang terstandarisasi dan memvalidasinya sehingga dapat digunakan untuk mengukur kemampuan berpikir analitis dan representasi kimia siswa pada topik laju reaksi. Soal yang dikembangkan mencakup keempat aspek kemampuan berpikir analitis dan keempat level multipel representasi (makroskopik, mikroskopik, simbolik, dan matematis). Sekelompok ahli telah mengonfirmasi validitas konstruk dan validitas muka dari soal yang telah dikembangkan, Test of Analytical Thinking based on Multiple Representation (TAT-MR) berjumlah 32 item. TAT-MR divalidasi terhadap 449 siswa sekolah menengah. Karakteristik TAT-MR dianalisis menggunakan model Rasch dengan pendekatan *Partial Credit Model-1 Parameter Logistic* (PCM-1PL). Hasil analisis menunjukkan bahwa terdapat 22 item dari TAT-MR yang memiliki nilai reliabilitas sangat baik sehingga dapat disimpulkan bahwa TAT-MR dapat digunakan sebagai instrumen yang baik untuk mengumpulkan data. Hasil penelitian ini menyarankan bahwa TAT-MR akan berguna untuk mengukur kemampuan berpikir analitis dan representasi kimia siswa pada pembelajaran kimia.

Kata Kunci: ***berpikir analitis, representasi kimia, model Rasch, laju reaksi, validasi***

INTRODUCTION

The biggest challenge of education field is the enhancement of the students' thinking skills. In the 21st century, the development of students' potential not only focuse on the great emphasis on fostering students to become academic, knowledgeable, and independent, but also to promote students become innovative, creative thinkers, effective doers, and skillful problem solvers. Hence, the development of higher order thinking skills is important for students to develop and promote a rational thinking. The characteristics of higher order thinking skills are identified as the analytical, evaluation, and create skills (Petrovska & Veseliovska, 2013).

An analytical thinking skill is considering as an extension of *understanding* of phenomenon and as a prelude to *evaluating* or *creating*. These thinking ability make the students to break the things (situations, practices, problems, statements, ideas, theory, arguments) into their component parts and establish how each parts are related each other and to an overall structure or objective (Anderson &Krathwohl, 2001; Thaneerananon, Triampo, &Nokkaew, 2016). The indicators of analytical thinking skill include cognitive process of differentiating (Anderson &Krathwohl, 2001; Areesophonpichet, 2013; Mayer, 2002; Ramirez &Ganaden, 2008; organizing (Anderson & Krathwohl, 2001; Areesophonpichet, 2013); and attributing (Areesophonpichet, 2013; Mayer, 2002; Ramirez &Ganaden, 2008). Analytical thinking leads the students to differentiate between facts and opinions, similarities and differences, and causes and effects. After that, the students also have to compare and analyze the consistent and contrary or irrational information given. Then, students have to identify the key matters by summarize that relevant information into one concept. Consequently, a learning which promote analytical thinking skill is a learning to determine the relevant or important parts of information, the ways in which the parts of an information are configured, and the underlying the purpose of those information.

Having analytical thinking is necessary to the students. It enhances meaningful learning because analytical reasoning is the basic way which used to solve a problem in the various field. Taleb & Chadwick (2016) identified chemistry learning which fosters analytical thinking skill means encouraging the students to analyze, criticize, assess, compare, and evaluate. Many students find any difficulties in applying these thinking skill to solve a common problem in the chemistry lesson because they have a lack of exercise to promote analytical thinking skill.

The analytical thinking skill makes the students understand the chemistry concept on its nature. The nature of chemistry can be viewed by the multiple representation level. The multiple representation level including the macroscopic, sub-microscopic, symbolic, and mathematics (Gilbert & Treagust, 2009; Hafsah, Hashim, Zurida, Jusoff, & Yin, 2014). The chemistry concept which provides with multiple representations is necessary to learn because of the difficulties in understanding the further concept if the initial concept not yet mastered. In fact, the representation which often raised by the teacher in developing a test tends to the symbolic and mathematics aspect only, while the macroscopic and sub-microscopic aspects are rarely to found. These facts lead to the lack of students' ability in solving the macroscopic and sub-microscopic aspects.

In the chemistry learning instruction, required a good students' understanding of the chemistry concept. The students' understanding of chemistry concept can be viewed by the multiple representation levels. Recent study showed that the use of multiple representation instruction leads the improving of students' concept understanding (Abdurrahman, Liliasari, Rusli, & Waldrip, 2011). The multiple representation level consisting of macroscopic, sub-microscopic, symbolic (e.g. Kozma, 2003; Treagust, Chittleborough, & Mamiala, 2003; Gilbert & Treagust, 2009) and mathematics (Hafsah, et al., 2014) can be used to describe the chemical concept. The first is the macroscopic level that represents the chemistry concept obtained by the experience or experiment (Li & Arshad, 2014). Johnstone (2000) proposed that the macroscopic level can be seen, touched, and felt. The second is the sub-microscopic level which identified as the chemical representation in the form of the visualization including the atom, ion, and molecule on the chemical reaction (Bucat & Macerino, 2009). Davidowitz, Chittleborough, & Murray (2009), state that the representation of the sub-microscopic level is expected to provide a complete description of the chemical reaction. The third is the symbolic level which represents the chemistry that consists the symbol or icon as a tool to describe the atom, characteristics, phase, and the equation of chemical reaction (Talanquer, 2011). This symbolic description including the writing of the element, compound, substance phase, graphics and table representation, and also writing the equal chemical equation. The last is the mathematical level which identified as the representation of the chemical calculation. These mathematical calculation leads the students to gain a better problem-solving ability on the problems of chemistry concept (Hafsah, et al., 2014).

Previous studies conclude that students have less competence in sub-microscopic representation (Kellya & Hansenb, 2017; Milenkoviæ, Segedinac, Hrin, & Cvjetiæanin, 2014). The students' inability to represent chemical phenomena at the sub-microscopic level can inhibit the ability to solve chemistry problems related to both macroscopic and symbolic phenomena (Chandrasegaran, Treagust, & Mocerino, 2007); Kozma, (2003). Other hand, the students were found difficulties in understanding the macroscopic representation into sub-microscopic and symbolic representations (Devetak, 2009; Davidowitz, et al., 2010). The difficulties faced by the students on the chemistry learning found when they combining the three level of chemical representations (Johnstone, 2000). In addition, Hafsahetal. (2014) proposedthat the mathematical ability affect the students' skill on solving the chemistry problems. A lack of mathematical ability leads the poor students' abilities in solving those chemistry problems. Consequently, the poor understanding of the chemistry concepts were caused by the chemistry learning which involve the multiple representation level which has not been emphasized (Sunyono, Yuanita, & Ibrahim, 2015). Hence, to assess students' representation on the chemistry concept should be measure by the test covering the fourth level of multiple representations.

One of chemistry topic that need to be understand through the fourth level of multiple representation is 'rate of reaction' topic. 'Rate of reaction' concept is largely abstract and needs to be supported by visualizing the abstraction in various representations to achieve conceptual understanding. The Rate of reaction topic involves in several concepts such as the concept of chemical reaction, collision theory, factors affecting rate of reaction, equation of rate of reaction and level of reaction. Macroscopic aspect can be directly observed from the experiment about the factors affecting rate of reaction. While, the sub-microscopic aspect could be learned by the collision theory and the application on the rate of reaction factors. Furthermore, the symbolic and mathematics aspects learned through the formula, table, and graph of the rate of reaction concept, enthalpy and activation energy, rate of reaction equation, and the reaction order. Hence, the teacher should assess the students'understanding in rate of reaction topics that covered these fourth level of multiple representation.

Recent studies on 'rate of reaction' topicwere concentrate on varied students' perspectives. Several researcher focused in overcoming students' alternative conception (Çakmakçý, Leach & Donnelly, 2006; Çalik, Kolomuç, & Karagolge, 2010; Kolomuç & Çalik, 2012); facilitating students' conceptual change (Kaya &Geban 2012; Kýrýk & Boz, 2012; Supasorn & Promarak, 2015); enhancing students' achievement (Redhana & Merta, 2017; Kurt & Ayas, 2012) and attitude toward chemistry (Seçken & Seyhan, 2015); improving students' critical thinking (Pratiwi, Rahayu, & Fajaroh, 2016); and also promoting students' anxiety (Olakanmi, 2015). Unfortunately, the students' has difficulties in learning 'rate of reaction' topic (Cakmakci, et al., 2006; Calik, Kolomuc & Karagolge, 2010; Taºtan, Yalcinkaya, & Boz, 2010; Turányi & Tóth, 2013). The studies afforementioned showed that none of them not integrate the analytical thinking and multiple representation level. It can be a reason that chemistry teacher difficulties to measure students' chemical representations comprehensively. The fourth level of multiple representation should be hand in hand in order to provide a comprehensive understanding of the students which brings the students stored their knowledge in their long-term memories. Thus, an instrument that covered the fourth level of multiple representation on the rate of reaction topic is need to be construct to promote students' analytical thinking.

The construction of a test based on multiple representation approaches is needed. It can use to explore deeply the students' ability of chemical representation. Moreover, the study also construct the test based on aspects of analytical thinking as the part of higher order thinking skills. Hence, the test is more useful because has two dimensions to explore the students' achievement. It will be the pilot study to construct alternative of assessment to support chemistry learning. Such a new assessment would first have to be demonstrated to be a valid and reliable measure of students' achievement.

The research purposes to constructof valid and reliable thetest of analytical thinking based on multiple representation (TAT-MR) of rate of reaction topic.

METHODS

Instrument Development

Instrument development was conducted according to Trochim work (1999) and respected the suggestion of Dalgety, Coll, & Jones (2003) related to the theoretical framework and concept of construct validity. To confirm content validity, the researchers elaborated a theoretical framework about analytical thinking and multiple representations level as a basis to construct the item and then ensured it to experts.

The analytical thinking framework was synthesized from several experts following Anderson & Krathwohl (2001), Mayer (2002), Ramirez & Ganaden (2008), Areesophonpichet (2013). The aspects of the analytical thinking construct as the result of these activities consisting of (1) differentiate, (2) organize, and (3) attributed. In addition, the instrument being developed of analytical thinking was based on the multiple representation frameworks. Besides ensuring content validity, a focus group discussion of experts (five chemistry educators) was used to gain face validity. Specifically, experts review all of the items for readability, clarity, and comprehensiveness and come to some level of agreement as to which items should be included in the final scale (Sangoseni, Hellman, & Hill, 2013). For enhancing the readability and usability, the TAT-MR was viewed by 10 high school chemistry teachers.

Sample for Validation

A total of 449 students of 11th grade from six high schools located in Yogyakarta, Indonesia were used as samples to get the data of the validity and reliability of the TAT-MR. The sample was gotten by purposive sampling, divide into two group criteria. First, high level school identified from the top 10 schools with the highest national exam results in Yogyakarta year 2017 and the second one as moderate high school that had a ranking of national exam among 15 - 30. There were three schools from each criterion that was gotten by randomized. A total of 243 students came from higher level school and 206 students from lower high school criteria.

Data Analysis

The analysis of the characteristics of the test instrument was conducted using Rasch model by Winstep program. Partial Credit Model 1-Parameter Logistic (PCM-1PL) approach was used for analysis. The analysis of the characteristics of the test instrument using Rasch model were assessed by the item fit analysis, reliability of person and item, the item difficulty, and information function with the standard error measurement. Before doing the analysis of the Rasch modelling, there were several assumptions could meet. The several assumptions in this case consisting the uni-dimensionality test, the local independency test, and the parameter invariance test (Hambleton & Swaminathan, 1985; Hambleton, Swaminathan, & Rogers, 1991). The uni-dimensionality test is the construct validity analysis which analyzed by the confirmatory factor analysis (CFA). In confirming whether the data is appropriate for factor analysis or not, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO-MSA), Barlett Sphericity test, and the anti-image correlation on varimax rotation were conducted.

RESULTS AND DISCUSSION

Results

The initial TAT-MR consisting of 32 open-ended questions that contain the indicator of analytical thinking and multiple representation aspects on each item. Five experts had reviewed the content and face validity of the initial TAT-MR. There was no item of TAT-MR should be added or reduced from this process. Hence, all of the items were used for further analysis. Table 1 presents the aspects of the three abilities on each item instrument developed.

The next step of the analysis consisting the construct validity and item characteristics analyses. The items characteristics analyses were conducted by Rasch model with PCM-1PL approach. The test assumptions that should be meet in using this model consisting the uni-dimensionality test (construct validity), the local independency test, and the parameter invariance test.

The Uni-dimensionality Assumption Test (Construct Validity)

The theuni-dimensionality test was conducted by the KMO-MSA and Barlett

Table 1. Indicators of TAT-MR on Aspects of Analytical Thinking and Multiple Representations

No.	Aspect of analytical thinking	Aspect of multiple representation
1.	Organize	Symbolic
2.	Attributed	Sub-microscopic
3.	Organize	Symbolic
4.	Attributed	Mathematics
5.	Differentiate	Sub-microscopic
6.	Organize	Symbolic
7.	Organize	Mathematics
8.	Differentiate	Macroscopic
9.	Attributed	Symbolic
10.	Organize	Sub-microscopic
11.	Differentiate	Symbolic
12.	Attributed	Symbolic
13.	Attributed	Mathematics
14.	Organize	Macroscopic
15.	Organize	Symbolic
16.	Differentiate	Symbolic

No.	Aspect of analytical thinking	Aspect of multiple representation
17.	Organize	Sub-microscopic
18.	Differentiate	Mathematics
19.	Differentiate	Symbolic
20.	Organize	Symbolic
21.	Differentiate	Symbolic
22.	Attributed	Mathematics
23.	Organize	Macroscopic
24.	Attributed	Symbolic
25.	Attributed	Sub-microscopic
26.	Organize	Symbolic
27.	Attributed	Mathematics
28.	Attributed	Sub-microscopic
29.	Differentiate	Macroscopic
30.	Organize	Symbolic
31.	Differentiate	Sub-microscopic
32.	Organize	Mathematics

Sphericity test. The results of these analysis presented in the Table 2.

Table 2. The Result of KMO-MSA and Barlett Sphericity Test

Test	Analytical Thinking	Conclusion for factor analysis
KMO-MSA test	.86	Appropriate
The significance value of Barlett Sphericity test*	.00	Appropriate

*statistical significance level of .05

According to Table 2, it can be seen that the KMO value of analytical thinking instrument was above .5 (.86 >.5) and it proves that the sample used is adequate. While the BarlettSphericity test shows that the variable among this study is correlate (.00 <.05). Hence, the data obtained in this study is appropriate for factor analysis on uni-dimensionality or construct validity investigation.

The other way to seek the unidimensionality assumption test is by the Scree plot. The scree plot analysis which clarified the eigen values is presents in the Figure 1.

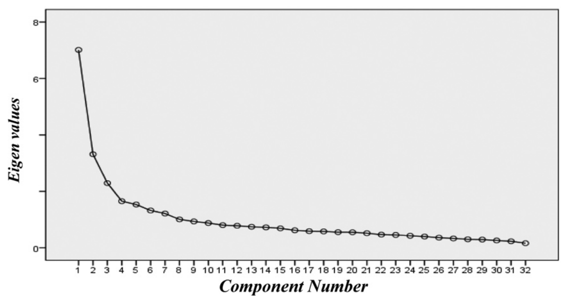

Figure 1. Scree plot of CFA on uni-dimensionality assumption test

The scree plot on Figure 1 is confirms that the instrument of analytical thinking consisiting of 8 factors. According to Figure 2, the curve starts to slop on the ninth factors. Hence, at least as many as 8 factors were formed with the first factor is the dominant factor. The percentage of the first factor was 21.921%.

The Local Independency Assumption Test

The results of the local independence assumption test which is the covariance matrix in this study observed on the Table 3.

Table 3 shows that the covariance values on analytical thinking are approaching .00. Hence, it can be concluded that the local independence assumption test is fulfilled.

The Parameter Invariance Test

The parameter invariance test in two ways, the items invariance parameter and students' ability invariance parameter. The results of the

Table 3. Covariance Matrix of Analytical Thinking Skill

Columns	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	.02									
C2	.00	.00								
C3	.00	.00	.00							
C4	.00	.00	.00	.00						
C5	.00	.00	.00	.00	.00					
C6	.00	.00	.00	.00	.00	.00				
C7	.00	.00	.00	.00	.00	.00	.00			
C8	.00	.00	.00	.00	.00	.00	.00	.00		
C9	.01	.00	.00	.00	.00	.00	.00	.00	.00	
C10	.01	.00	.00	.00	.00	.00	.00	.00	.01	.02

items parameter invariance in this study can be seen in Figure 2.

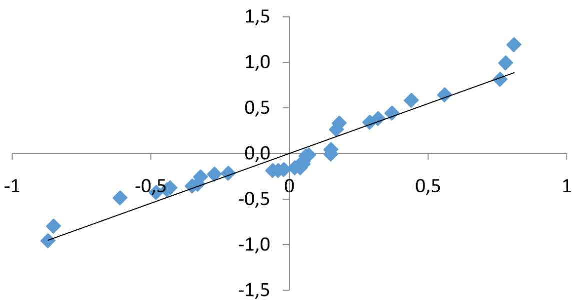

Figure 2. Scree Plot of Items Invariance Parameter

It can be concludes that both groups has a high correlation. Hence, it can be said that the items of the instrument are fulfilled the items parameter invariance assumption.

On the other hand, the students' ability invariance parameter was used to compare the students' ability on performing the items. The result of the students' ability invariance parameter showed in the Figure 3.

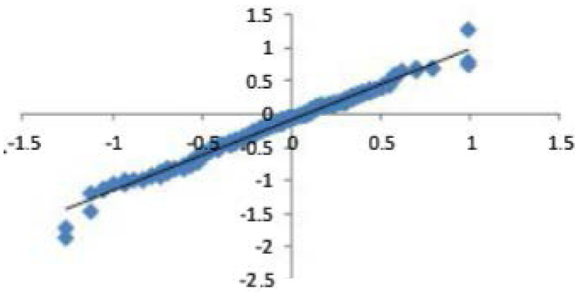

Figure 3. Scree Plot of Students' Ability Invariance Parameter

Figure 3 showed thatthe dot isform a linear line. Thus, it can be concluded that the students' ability parameter invariance assumption was fulfilled.

All of the assumptions of the Rasch modelling consisting the uni-dimensionality test, local independency test, and the invariance parameter test were fulfilled. Hence, the items characteristics analysis can be conducted by the Rasch model.

The Results of Items Characteristics Analysis

The characteristics of test instrument wereconducted following (a) itemfit model; (b) test reliability; (c) item difficulty; and (d) test information function.

Analysis of Instrument Item Fit

The result of the instrument fit analysis in this study presents on the Table 4.

Table 4 shows that from as many as 32 items which constructed, there were 10items that not fit with the model. These eight items are doesn't fit with the PCM-1PL model, hence the eight items consisting of number 1, 4, 5, 6, 7, 9, 12, 13, 22, and 32 were discarded.

The Reliability of the Instrument

The reliability analysis in this study was conducted by means of the Rasch measurement modelwhich is the classical test theory in the form of Cronbach's Alpha coefficients. There were two types of reliability in the Rasch model, the reliability of the Person and the Item measured. These two types of reliability were used two indicators of the person separation and item separation.The summary of the statistical

Table 4. Item Fit of Measurement Instrument

Item	Outfit MNSQ	Outfit ZSTD	Pt-Measure Correlation	Conclusion
1	.5581	-7.9994	.30	Not fit model
2	.5269	-9.0495	.40	Fit Model
3	.8952	-1.4691	.30	Fit Model
4	1.3157	2.9613	.30	Not fit model
5	.6628	-5.3893	.30	Not fit model
6	.5953	-7.5094	.30	Not fit model
7	1.3412	5.3713	.20	Not fit model
8	.7064	-4.4793	.40	Fit Model
9	1.6094	8.7516	.30	Not fit model
10	1.0867	1.1511	.40	Fit Model
11	1.1296	1.6111	.30	Fit Model
12	1.3412	4.8113	.20	Not fit model
13	2.1184	9.9021	.20	Not fit model
14	1.4793	7.2215	.50	Fit Model
15	.5277	-9.3895	.60	Fit Model
16	1.2588	3.2313	.40	Fit Model
17	.8732	-2.3391	.60	Fit Model
18	.7409	-2.9293	.50	Fit Model
19	.7300	-5.2293	.50	Fit Model
20	.9741	-.4290	.50	Fit Model
21	1.2083	3.4612	.60	Fit Model
22	1.6683	9.9017	.50	Not fit model
23	.7757	-4.0692	.60	Fit Model
24	.9452	-.8191	.60	Fit Model
25	.8480	-2.8492	.70	Fit Model
26	1.4640	7.0315	.50	Fit Model
27	.6192	-4.2494	.50	Fit Model
28	.8499	-1.8992	.50	Fit Model
29	.8103	-3.2392	.50	Fit Model
30	.8271	-2.9492	.50	Fit Model
31	.8274	-2.4092	.50	Fit Model
32	1.2635	2.2313	.30	Not fit model

Table 5. Summary Statistics of Person and Item Reliability

	Infit		Outfit				
Parameter (N)	**MNSQ**	**ZSTD**	**MNSQ**	**ZSTD**	**Separation**	**Reliability**	**Category**
Persons (449)	1.02	.0	1.02	.0	2.34	.85	High
Items (32)	1.01	-.7	1.02	-.3	12.78	.99	Excellent

result of the measurement instrument presents in the Table 5.

It can be seen from Table 6 that the personal and item reliability coefficients found to be .85 and .99 respectively. Hence, it can be concluded that the persons' reliability found in high category while the items reliability found in an excellent category.

Item Difficulty

The item difficulty in this study was obtained by Winstep program on the item measure value. The result of these analysis presents on the Table 6.

Based on Table 5, the index difficulty items are well distributed on the very easy, easy, medium, difficult, and very difficult category with the range of the index difficulty between -1.03 and 1.06.

Table 6. Item Difficulty of TAT-MR

Item Number	Difficulty index	Category
4*	-1.03	Very Easy
11	-.68	Easy
10	-.61	Easy
8	-.58	Easy
3	-.57	Easy
5*	-.54	Easy
1*	-.44	Medium
2	-.38	Medium
6*	-.37	Medium
9*	-.25	Medium
14	-.21	Medium
26	-.21	Medium
7*	-.2	Medium
21	-.15	Medium
22*	-.13	Medium
25	-.03	Medium
17	.02	Medium
19	.08	Medium
13*	.1	Medium
20	.11	Medium
23	.18	Medium
15	.23	Medium
30	.25	Medium
29	.26	Medium
12*	.31	Medium
24	.35	Medium
16	.5	Medium
31	.51	Difficult
28	.62	Difficult
18	.84	Difficult
27	.94	Difficult
32*	1.06	Very Difficult

*not fit model

Informational Function and Standard Error Measurement

The relationship between the IF with the SEM value in the analytical thinking skill presents in the Figure 5.

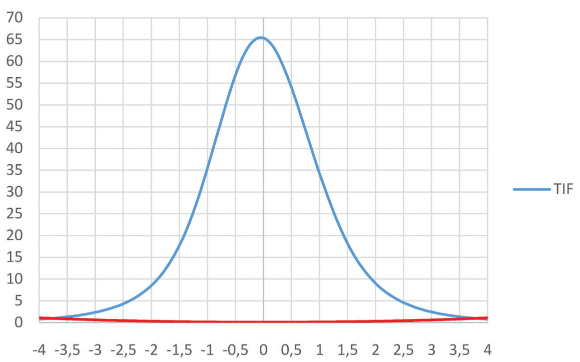

Figure 5. The Relationship of IF and SEM of Analytical Thinking Skill

Figure 5 shows that the maximum IF value of analytical thinking ability from a test instrument based on multiple representations with 32 items is found to be 65.46 at è about 0 logit and SEM of.12.

Discussion

The construction of the TAT-MR in this study was well organized. The amount of the items which being developed are 32 items, consisting the indicator of analytical thinking skill and the aspect of multiple representations. Content, face, construct validation, and the item characteristics analysis had been done to analyze the quality of the instrument.The TAT-MR has a high content and face validity declared by a group of expert. It can be seen from the results of the experts' response toward the TAT-MR that there was no item should be added or discarded in this instrument. Based on the experts' responses and comments, the instrument was revised and modified. The experts' feedback related to the missing in presenting sub-microscopic level in some concepts and related to the grammatically error, also the choice of an appropriate word. All of the feedbacks from the experts have been analyzed by the researcher and necessary revision has made. In conclusion, all of the items on TAT-MR has a good content and face validity.

In addition, the instrument is using the polytomous scales, hence the item characteristics analysiswassuitable usingtheRasch model withPCM-1PL approach. Several test assumptions were calculated before using the Rasch model with PCM-1PL. The first was the uni-dimensionality test that aims to test whether each item of the instrument is measure one variable or one ability only (Reckase, 1979). The uni-dimensionality assumption test also well-known as the construct validity of the instrument. If the uni-dimensionality test was fulfilled, so that the construct validity also fulfilled. The factor analysis was used to obtain the uni-dimensionality assumption test of the instrument. The objective of the factor analysis is to identify the relationship among variables by seeking the computational result on Eigen value in the matrix of intercultural variance-covariance.

In this study, the uni-dimensionality assumption test was initiate by the KMO-MSA test and BarlettSphericity Test, to know the data obtained in this study were appropriate for factor analysis or not. The KMO-MSA test aims to determine sample adequacy, while the BarlettSphericity test was used to determine whether any relationship or not among the variables. The factor analysis can't be conducted if the KMO-MSA value is less than the critical value which is .5 (Leech, Barret, & Morgan, 2005) and the Barlett Sphericity test >.05 (Beavers, Lounsbury,Richard, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S.., 2013). As seen from Table 2, the result of the KMO value found to be .86 and it proves that the sample used is adequate (.86 >.5). While the BarlettSphericity test shows that the variable among this study is correlate (.00 <.05). Hence, the data obtained in this study is appropriate for factor analysis on uni-dimensionality or construct validity investigation.

The construct validity aims to determine the items of the instrument are valid or not according to the empirical data. The construct validity was conducted by the interpretation of the anti-image value on the result of factor analysis. The anti-image correlation value obtained after the KMO-MSA and BarlettSphericity test were fulfilled. The factor analysis in proving the construct validity with the anti-image correlation value criteria above .05 can be concludes that the items are valid (Wahyuningsih, 2009). The result of the anti-image correlation in this study has value greater than .5 for each of the 32 items. Thus, the values of these items has a high contribution toward the factor structure of the instrument.

Another way to found the unidimensionality assumption test is by the Scree plot. The scree plot is used to describe the ilustration of the eigen values by the number of component preserve the factors. The unidimensionality test could be considered fulfilled if the instrument has dominant component which measures the ability being tested (Hambleton, et al., 1991; Guler, Uyanik, & Teker, 2014). Brown, Obasi, & Barret (2016) proposed that if there is factor dominant with the cumulative percentage greater than 20%, thus the uni-dimensionality assumption test can be considered fulfilled. As seen on the Figure 1, the output of the CFA in this study is generated by the first factor and it is able to describe the varinace greater than 20%. At least as many as 8 factors were formed with the first factor is the dominant factor (the percentage of the first factor found to be 21.921%). Hence, the unidimensionality assumption test of the instrument developed is fulfilled (Reckase, 1979).\

The second was the local independency assumption test aims to prove that the participant answer toward one items is not affect their answer toward the other items of the instrument. The local independency assumption test shows that if the ability which affect the performance test is constant, thus the participants' response toward each items of the instrument doesn't correlate statistically each other. Consequently, if the uni-dimensionality assumption test is accepted, the local independency assumption test also accepted (Retnawati, 2014).

In this study, the local independency assumption test was conducted by calculating the covariance matrix based on the students' ability on each instrument tested (Greiff,Wüstenberg, Molnár, Fischer, Funke, &Csapó, 2013). The local independency assumption test declared fulfilled if the value under the diagonal line on the variance-covariance matrix is .00. The .00 values indicate that the students' skill in answering the items is not affecting on their answering skill toward the other items of the instrument. In short, each items of the instrument are independence. According to Table 3, shows that the covariance values on students' analytical thinking and multiple representation are approaching .00. Thus, the local independence assumption test is fulfilled. The result of this study confirmed the idea proposed by Hambleton & Swaminathan (1985) that if the covariance value is approaching .00, consequently the local independency assumptions test is fulfilled.

The last test assumption was the parameter invariance assumption test that aims to prove the parameter invariance of items and participants' ability (Köse, 2014). The items parameter invariance test was conducted to determine the items characteristics consistency that answered by the different group of the students. Besides that, the students' ability parameter invariance was conducted to estimate the unchanging ability even the items are change. In the items invariance parameter, the participants are divided into two groups. The first group is the participant with the odd number, while the second group is the participant with the even number. Both groups are answering the same of items instrument. The items parameter invariance test conducted by calculating the difficulties level for both groups. These difficulties level of both groups is analyze by seeking the linear line.

According to Figure 2, in the term of items invariance parameter, it can be concluded that both groups has a high correlation. Almost every dot is on or approaching the linear line. Hence, it can be said that the items of the instrument are fulfilled the items parameter invariance assumption. Consequently, the items which developed are not affecting by the participants. Even the participants are different, the items characteristics is unchanging.In addition, the students' ability invariance parameter was used to compare the students' ability on performing the items. Figure 3 showed that the dot is form a linear line. Thus, it can be concluded that the students' ability parameter invariance assumption was fulfilled. In short, the students' ability is not affect by the items of the instrument.

All of the assumptions of the Rasch modelling consisting the uni-dimensionality test, local independency test, and the invariance parameter test were fulfilled. Hence, the items characteristics analysis can be conducted by the Rasch model with PCM-1PL approach. By this approach, the instrument characteristics analysis was conducted following (a) each item and each test participant's fit to the model; (b) the test has consistency in measurement (reliability); (c) each item estimating item difficulty level ranges between -2 logit d" bi d" 2 logit (Hambleton &Swaminathan, 1991); and (d) the test will provide good information if TIF e" 10 (Wiberg, 2004).

The item fit analysis was used to determine that the items have functioning normally in measuring or not. If an item of the instrument is not fit with the model, it can be said that there is any students' misconception toward the item. El-Korashy (1995) proposed that if the item is statistical fit with the model hence it is considered as the valid item. In addition, Boone, Staver, & Yale (2014) stated that there were three criteria of the item fit consisting of the value of output mean square (MNSQ) is accepted if .5< MNSQ <1; the value of Z-standard outfit (ZSTD) is accepted if -2.0< ZSTD <+2.0; the value of correlation points (Pt Mean Corr) is accepted if .4< Pt Mean Corr <.85. An item considered fit with the model if at least there were two criteria of item fit is accepted. The item fit analysis was done using Winstep program. The result of the analysis showed that overall, as many as 22 items is fit with the PCM-1PL model and can be used for the analysis of measuring students' analytical thinking skill and chemical multiple representations. In short, there are 22 items of the instrument is having good construct validity.

The instrument reliability was analyzed according to the person and the item analysis. According to Table 6, it can be seen that the person reliability coefficients found to be .85 while the item reliability coefficients was .99. The coefficient of person reliability of the instrument indicates that there was 85% consistency of the students' response toward all of the items in the instrument. While the item reliability coefficient implied that there was 99% certainty of the consistency test items in obtaining same result repeatedly. George & Mallery (2003) provide the category of the reliability coefficients consisting the excellent category, if the reliability coefficients > .90; good category if >.80; acceptable category if>.70; questionable category if > .60; poor category if >.50; and unacceptable category if <.05. Hence, the personal reliability shows a good category, while the items reliability showed an excellent category. Hence the intrument in this study have a high consistency to measure the students' analytical thinking and multiple representation ability with a very minimum error (Retnawati, 2016).

The next items characteristics analysis was the index difficulty that provides to find out the correct answer opportunity of a problem at certain of ability level. The parameter of the item difficulty is expressed in logit units. A good instrument item has a range of item difficulty between -2.0 logit and +2.0 logit (Hambleton & Swaminathan, 1985). An item considered as a too difficult item if they have index difficulty above +2.00 logit while if they have index difficulty under -2.0 logit it is considered as too easy item.

This study refer on the interpretation of difficulty value following Adedoyin & Mokobi (2013) who states that an item categorized very difficult if the value of b (item measure) e" +1; difficult + .5 d" b < +1; medium -.5 d" b < + .5; easy 1 d" b < - .5; and very easy b d" 1. Hence based on Table 5, the resultsof index difficulties items are well distributed on the very

Item number 32
The reaction which occurred on an apple was continued by varying the vitamin C concentration and the apple extract concentration. Assume the vitamin C is "X" and the apple extracts is "Y". The result of the experimental observation presents below.

No	X concentration (M)	Y concentration (M)	Time (Second)
1	.01	.1	864
2	.02	.4	54
3	.03	.3	32
4	.04	.2	27

According to the data above, determine the value of K (constant)!

easy, easy, medium, difficult, and very difficult category with the range of the index difficulty between -1.03 and 1.06. According to these range value, it can be said that the item of the instrument has a good difficulty index. A well distributed of the item difficulty index lead the instrument considered as a good tool to measure the students' content understanding from low to high level ability. A very difficult items was found on the item number 32. This item consists of mathematics multiple representation with the indicator analytical thinking of organize. The item number of 32 presents in Figure 4.

*Item number 32*The reaction which occurred on an apple was continued by varying the vitamin C concentration and the apple extract concentration. Assume the vitamin C is "X" and the apple extracts is "Y". The result of the experimental observation presents below.**NoX concentration (M)Y concentration (M)Time (Second)**1.01.18642.02.4543.03.3324.04.227

According to the data above, determine the value of K (constant)!

Based on the illustration of item with very difficult criteria present above, majority of the students find difficulties to solve that problem. The students feel that the reaction order is difficult to determine because they used to solve the problem with the same concentration on each variable of observation data. On this problem, there were no concentration which same on each variable, hence the students feel confuse on determining the reaction order. Just in case if the students can organize the number on the table in a good manner with the quadratic equation, thus the problem become very easy.

The last analysis was the informational function (IF) test. IF test was used to further description about the reliability coefficient of the overall test of items. The test will provide a good information if TIF e" 10 (Wiberg, 2004). After the information function is estimated, then the standard error of measurement (SEM) calculated. In item response theory, measurement error is closely related toward the information function. The greater the value of the IF, the smaller the SEM value or vice versa (Hambleton &Swaminathan, 1991). Figure 5 shows that the maximum IF value of analytical thinking ability from a test instrument based on multiple representations with 32 items is found to be 65.46 at è about 0 logit and SEM of .12. These results indicate that the participants in this study who conducted the tests of the instrument based analytical thinking and multiple representation ability provide good information with the smallest measurement error done by the students who have the ability about 0 logit.

CONCLUSION

Development of a good instrument is a long process by iterative procedure. The final version of TAT-MR has 22 items that fit with the Rash model. It is consisting of three aspects of analytical thinking skills and four levels of MR. The number of item on each analytical thinking skills aspects were well distributed on 8 items ofdifferentiateaspect, 9 items on organize aspect, and 5 items covering on attributed aspect. Meanwhile, the level of MR consisting the macroscopic level of 4 items, the sub-microscopic level spreading on 6 items, the symbolic level covering the 10 items, and also the mathematics level with 2 items.The reliability of person and item found in high and excellent category respectively. The item difficulty was well distributed on very easy, easy, medium, difficult,until very difficult category. According to IF test, it is indicating that the participants in this study who conducted the tests provide good information with the smallest measurement error done by the students who have the ability about 0 logit.Moreover, the result of this study showedthat the TAT-MR have a high validity in content, face and construct. This suggests that the TAT-MR is potential to be a useful instrument for chemistry teachers and researcher for measuring the students' analytical thinking and students' chemical representation in rate of reaction topic.

ACKNOWLEDGEMENT

Thanks to the reviewers for the suggestions and feedback.

REFERENCES

Abdurrahman, Liliasari, Rusli, A., & Waldrip, B. (2011). Implementasi pembelajaran berbasis multi representasi untuk peningkatan penguasaan konsep fisika kuantum. *Cakrawala Pendidikan, 30*(1), 30-45. doi:10.21831/cp.v1i1.4189.

Adedoyin, O. O., & Mokobi, T. (2013). Using IRT psychometric analysis in examining the quality of junior certificate mathematics multiple choice examination test items. *International Journal of Asian Social Science, 3*(4), 992-1011.

Anderson, L. W., &Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing.* New York, NY: Longman.

Areesophonpichet, S. (2013). A development of analytical thinking skills of graduate students. *The Asian Conference on Education* (pp. 1-5). Osaka, Japan: The International Academic Forum.

Beavers, A. S., Lounsbury, J. W., Richard, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation, 18*(6), 1-13.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences.* Dordrecht: Netherlands.

Bucat, B., &Mocerino, M. (2009). Learning at the sub-micro level: structural representations. In Gilbert, J. K., &Treagust, D. F. (Eds), *Multiple representations in chemical education.* (pp.1-8). Dordrecht: Springer.

Brown, R. L., Obasi, C. N., & Barret, B. P. (2016). Rasch analysis of the WURSS-21 dimensional validation and assessment of invariance.*J. Lung. Pulm. Respir. Res., 3*(2), 1-16. doi:10.15406/jlprr.2016.03.00076.

Çakmakçý, G., Leach, J., & Donnelly, J. (2006). Students' ideas about reaction rate and its relationship with concentration or pressure. *International Journal of Science Education, 28*(15), 1795-1815. doi:10.1080/09500690600823490.

Calik, M., Kolomuc¸ A., & Karagolge, Z. (2010). The effect of conceptual change pedagogy on students' conceptions of rate of reaction. *Journal of Science Education and Technology*, *19*(5), 422-433. doi:10.1007/s10956-010-9208-9.

Chandrasegaran, A. L., Treagust, D. F., &Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice, 8*, 293-307. doi:10.1039/b7rp90006f.

Dalgety, J., Coll, R. K., & Jones, A. (2003). Development of chemistry attitudes and experiences questionnaire (CAEQ). *Journal of Research in Science Teaching, 40*(7), 649-668. doi: 10.1002/tea.10103.

Davidowitz, B., Chittleborough, G., & Murray, E. (2010). Student-generated submicro diagrams: a useful tool for teaching and learning chemical equations and stoichiometry. *Chem. Educ. Res. Pract., 11*(3), 154-164.doi: 10.1039/C005464J.

Devetak, I. E. (2009). Comparing Slovenian year 8 and year 9 elementary school pupils' knowledge of electrolyte chemistry and their intrinsic motivation. *Chem. Educ. Res. Pract., 10*(1), 281-290.doi: 10.1039/B920833J.

El-Korashy, A. F. (1995). Applying the rasch model to the selection of items for a mental ability test. *Educational and Psychological Mesurement, 55*(5), 753-763. doi:10.1177%2F0013164495055005006.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference. 11.0 update (4th Ed.).* Boston, MA: Allyn & Bacon.

Gilbert, J. K., & Treagust, D. F. (2009). Macro, submicro and symbolic representations and the relationships between them: Key models in chemical education. In J. K. Gilbert & D. F. Treagust (Eds), *Multiple representations in chemical education.* (pp.1-8). Dordrecht: Springer.

Guler, N., Uyanik, G. K., & Teker, G. T. (2014). Comparison of classical test theory and item response theory in terms of item parameters. *European Journal of Research on Education, 2*(1), 1-6.

Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., &Csapó, B. (2013). Complex problem solving in educational contexts—something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology, 18*(1), 1-15. doi:10.1037/a0031856.

Hafsah, T., Hashim, R., Zurida, I., Jusoff, K., & Yin, K. Y. (2014). The influence of students' concept of mole, problem representation ability and mathematical ability on stoichiometry problem solving. *Scottish Journal of Arts, Social Sciences and Scientific Studies, 21*(1), 3-21.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory principles and applications*. Boston, MA: Kluwer Nijhoff Publishing.

Hambleton, R. K., Swaminathan, H., &Rogers, H. J. (1991). *Fundamental of item response theory*. Los Angeles: Sage Publication, Inc.

Johnstone, A. (2000). Chemical education research: Where from here? *Chemistry Education, 4*(1), 34-48.

Kaya, E. & Geban, O. (2012). Facilitating conceptual change in rate of reaction concepts using conceptual change oriented instruction. *Education and Science, 37*(163), 216-225.

Kellya, R., & Hansenb, S. (2017). Exploring the design and use of molecular animations that conflict for understanding chemical reactions. *ACS symposium on Chemical Education, 40*(4), 476-481. doi:10.21577/0100-4042.20170043.

Kýrýk, O. T &Boz, Y. (2012) Cooperative learning instruction for conceptual change in the concepts of chemical kinetics. *Chemistry Education Research and Practice, 13*(3), 221–236.doi:10.1039/C1RP90072B.

Kolomuç, A. & Çalýk, M. (2012) A comparison of chemistry teachers'and grade 11 students' alternative conceptions of 'rate of reaction'. *Journal of Baltic Science Education, 11*(4), 333-346.

Köse, I. A. (2014). Assessing model data fit of unidimensional item response theory models in simulated data. *Academic Journal Educational Research and Reviews, 9*(17), 642-649. doi:10.5897/ERR2014.1729.

Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction, 13*(2), 205-226.doi:10.1016/S0959-4752(02)00021-X.

Kurt, S., & Ayas, A. (2012). Improving students' understanding and explaining real life problems on concepts of reaction rate by using a four step constructivist approach.

Energy Education Science and Technology Part B: Social and Educational Studies, 4(2), 979-992.

Leech, N. L., Barret, K. C., & Morgan, G. A. (2005). *SPSS for intermediate statistics: Use and interpretation.* New Jersey, NJ: Lawrence Erlbaum Associates, Inc.

Li, W. S., & Arshad, M. Y. (2014). Application of multiple representation levels in redox reactions among tenth grade chemistry teachers. *Journal of Turkish Science Education, 11*(3), 35-52.doi:10.12973/tused.10117a.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice, 41*(4), 226-232. doi:10.1207/s15430421tip4104_4.

Milenkoviæ, D., Segedinac, M., Hrin, T., & Cvjetiæanin, S. (2014). Cognitive load at different levels. *Croatian Journal of Education, 16*(3), 699-722.

Olakanmi, E. (2015). The effects of a web-based computer simulation on students' conceptual understanding rate of reaction and attitude towards chemistry. *Journal of Baltic Science Education, 14*(5), 627-640.

Petrovska, S., & Veselinovska, S. S. (2013). Contemporary pedagogical approaches for developing higher level thinking on science classes. *Procedia – Social and Behavioral Sciences, 92,* 702-710. doi:10.1016/j.sbspro.2013.08.72.

Pratiwi, Y., Rahayu, S., & Fajaroh, F. (2016). Socioscientific issues (SSI) in reaction rates topic and its effect on the critical thinking skills of high school student. *Jurnal Pendidikan IPA Indonesia, 5*(2), 164-170. doi:10.15294/jpii.v5i2.7676.

Ramirez, R. P. B., & Ganaden, M. S. (2008). Creative activities and students' higher order thinking skills. *Education Quarterly, 66*(1), 22-33. doi: 10.1.1.824.9279

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*(3), 207-230. doi:10.3102/10769986004003207.

Redhana, I. W., & Merta, L. M. (2017). Green chemistry practicum to improve student learning outcomes of reaction rate topic. *Cakrawala Pendidikan, 34*(3), 382-403. doi:10.21831/cp.v36i3.13062.

Retnawati, H. (2014). *Teori respon butir dan penerapannya untuk peneliti, praktisi pengukuran, dan pengujian mahasiswa pascasarjana.* Yogyakarta: Parama Publishing.

Retnawati, H. (2016). *Validitas reliabilitas & karakteristik butir*. Yogyakarta: Parama Publishing.

Sangoseni, O., Hellman, M., & Hill, C. (2013). Development and validation of a questionnaire to assess the effect of online learning on behaviors, attitudes, and clinical practices of physical therapist in the United States regarding evidence-base clinical practice. *The International Journal of Allied Health Sciences and Practice, 11*(2), 1-12.

Seçken, N & Seyhan, H.G. (2015). An analysis of high school students' academic achievement and anxiety over graphical chemistry problems about the rate of reaction: The case of Sivas province. *Procedia-Social and Behavioral Sciences, 174,* 347-354.doi:10.1016/j.sbspro.2015.01.671.

Sunyono, Yuanita, L., & Ibrahim, M. (2015). Mental models of students on stoichiometry consept in learning by method based on multiple representation. *The Online Journal of New Horizons in Education, 5*(2), 30-45.

Supasorn, S., & Promarak, V. (2015). Implementation of 5E inquiry incorporated with analogy learning approach to enhance conceptual understanding of chemical reaction rate for grade 11 students. *Chemistry Education Research and Practice, 16*(1), 121–132.doi:10.1039/c4rp00190g.

Taleb, D. M., & Chadwick, C. (2016). Enhancing student critical and analytical thinking skills at a higher education level in developing countries: Case study of the british university in dubai. *Journal of Educational and Instructional Studies, 6*(1), 67-77.

Talanquer, V. (2011). Macro, submicro, and symbolic: The many faces of the chemistry "triplet". *International Journal of Science Education, 33*(2), 179-195. doi:10.1080/09500690903386435.

Tastan, Ö., Yalcinkaya, E., & Boz, Y. (2010). Pre-service chemistry teachers' ideas about reaction mechanism. *Journal of Turkish Science Education, 7*(1), 47-60.

Thaneerananon, T., Triampo, W., & Nokkaew, A. (2016). Development of a test to evaluate students' analytical thinking based on fact versus opinion differentiation. *International Journal of Instruction, 9*(2), 123-138. doi:10.12973/iji.2016.929a.

Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2003). The role of subsub-microscopic and symbolic representations in chemical explanations. *International Journal of Science Education, 25*(11), 1353-1368. doi:10.1080/09500690320000070306.

Trochim, W.M. (1999). The *research methods knowledge base* (2nd Ed.). Cincinnati, OH: Atomic Dog.

Turányi, T., & Tóth, Z. (2013). Hungarian university students' misunderstandings in thermodynamics and chemical kinetics. *Chem. Educ. Res. Pract., 14*(1), 105–116. doi:10.1039/c2rp20015e.

Wahyuningsih, H. (2009). Validitas konstruk alat ukur spirituality orientation inventory (SQI). *Jurnal Psikologi, 36*(2), 116-129. doi:10.22146/jpsi.7890.

Wiberg, M. (2004). *Classical test theory vs item response theory: An evaluation of the theory test in the Swedish driving-license test.* Santiago: Centro de EstudiosPúblicos.