

The characteristics of chemistry test items on nationally standardized school examination in Yogyakarta City

Ummul Karimah^{1*}; Heri Retnawati¹; Deni Hadiana²; Pujiastuti³; Eri Yusron¹

¹Universitas Negeri Yogyakarta

Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia

²Center for Assessment and Learning of the Ministry of Education and Culture

Jl. Gunung Sahari Eks Komp. Siliwangi No. 4, Pasar Baru, Sawah Besar, Jakarta 10710, Indonesia

³Sekolah Menengah Atas Negeri 7 Yogyakarta

Jl. MT. Haryono No. 47, Suryodiningratan, Mantrijeron, Yogyakarta 55141, Indonesia

*Corresponding Author. E-mail: ummulkarimah01234@gmail.com

ARTICLE INFO

Article History

Submitted:

22 April 2020

Revised:

8 November 2020

Accepted:

19 December 2020

Keywords

difficulty index;
discrimination index;
reliability

Scan Me:



ABSTRACT

This study aimed to describe the characteristics of chemistry test items in Nationally Standardized School Examination or *Ujian Sekolah Berstandar Nasional* (USBN), consisted of discrimination index, difficulty index, and question reliability. The data were collected by using documentation of the answers of 194 students. The type of research was descriptive exploratory with quantitative and qualitative approaches. Quantitative data analysis was performed by using the classical test theory approach and item response theory with 1 logistical parameter. Meanwhile, qualitative data analysis was conducted to describe the items categorized as difficult and bad categories. The results show that, according to classical test theory, the USBN Chemistry item test has an average level of difficulty of 0.57, which is categorized as moderate. Regarding the discrimination index, the average of different test power obtained is equal to 0.146 and belongs to the good category. Based on the item response theory, the average difficulty index of -0.00086 is categorized as moderate. The results of the estimated reliability of the questions amounted to 0.48 is included in the moderate category. The results of the qualitative analysis show that the items which belong to the difficult category are the items on salt hydrolysis material, acid-base titration, the concept of periodicity, manufacture, and use of chemical compounds according to classical test theory. Meanwhile, according to the theory of response, items that belong to the difficult category were found in acid-base titration material, the colligative solution phenomenon, and the concept of periodicity of the main metal elements.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



How to cite:

Karimah, U., Retnawati, H., Hadiana, D., Pujiastuti, P., & Yusron, E. (2021). The characteristics of chemistry test items on nationally-standardized school examination in Yogyakarta City. *REiD (Research and Evaluation in Education)*, 7(1), 1-12. doi:<https://doi.org/10.21831/reid.v7i1.31297>

INTRODUCTION

Assessment is one of indicators to measure the success of a testing process. A good assessment makes a good education quality. The results of the assessment can determine the success of the future education process (Mardapi, 2016, p.10). One of the educational processes occurs in the learning process in classroom. Teacher is one of the factors that determine students' learning success (Akbar, 2016, p.6). Competent, skilled, and highly dedicated teachers can enhance students' achievement in the classroom.

However, in the learning process, one of the difficulties is caused by chemistry subjects that are less liked by students. The cause of this difficulty is that chemistry subjects study a lot of abstract things. According to Gabel (Ristiyani & Bahriah, 2016, p.19), this abstract makes chemis-

try a complex subject. Coll and Taylor (Ristiyanı & Bahriah, 2016, p.19) added that there was difficulty in understanding chemical concepts due to the inability to connect the macroscopic and microscopic world or also called terms of triangle as macroscopic, submicroscopic, and symbolic (Taber, 2013, p.156). These concepts consisted of concept of mole, atomic structure, kinetic theory, thermodynamics, electrochemistry, chemical change and reactivity, equalization of redox reaction equations, and stereochemistry.

One example of test which aims to determine students' ability in Chemistry at the third year in high school level is the implementation of USBN. Regarding to Minister of Education and Culture Regulation Number 4 Article 1 Paragraph 5 of 2018, USBN is a form of test which aims to measure the achievement of students' competencies conducted by certain education units that refer to Graduates Competency Standards in order to get recognition of learning achievements, especially in chemistry subjects. USBN question scripts consist of 20% to 25% items given by the Ministry and 75% to 80% are given by the teachers and discussed in Teacher Working Group (*Kelompok Kerja Guru* or KKG), Teacher Subject Meeting (*Musyawarah Guru Mata Pelajaran* or MGMP), Tutors Forum, and *Salafiah* Islamic Boarding School Teacher Working Group (*Kelompok Kerja* or Pokja PPS) regulated in the same regulation in Article 11 Paragraph 1. This shows that the role of teachers is important in the preparation of an instrument for evaluating the implementation of USBN as a standard for the graduation of students in schools.

The preparations for the implementation of the USBN seem relatively short by the sudden distribution of the question grid by the ministry which causes the teachers who make the questions only have a short amount of time. Besides, the teachers who do not really master the preparation of good assessment instruments (Retnawati, 2015, p.400) is one of the determinants of the quality of USBN questions, especially in Yogyakarta City.

A good assessment quality is a matter that fulfills the requirements of good quality characteristics. The characteristics of this item can be seen based on the level of difficulty, discrimination index, reliability, and measurement error parameters (Herkusumo, 2011, p.457). According to classical test theory (Herkusumo, 2011, p.457), the level of difficulty of an item explains the percentage of students' correct answers that are divided into easy, moderate, and difficult categories. The power of differentiation shows the ability of an item in order to distinguish intelligent students from low student ability. It is often assumed that items that have a low coefficient of differentiation interpret that high student ability cannot answer the question compared to low students ability. This makes the items become bad items. The item reliability indicates the level of trustworthiness of an assessment instrument even though it has been used repeatedly. The incongruity of an instrument certainly makes the quality of the assessment become not good.

This study is relevant to the research conducted by Ariani et al. (2019, p.42) who also analyzed USBN questions but in Mathematic subject that were centred on analyzing student errors in solving USBN problems. These errors consisted of errors in reading, understanding, transformation, process skills, and also writing answers. The item analysis research was also conducted by Yustika et al. (2015, p.1330) who analyzed the characteristics of chemistry subject items but in the school examination or *Ujian Akhir Sekolah* (UAS) question.

Based on the background of the problem and the study of the previous studies, it can be concluded that the analysis of item characteristics, both using classical test theory and item response theory is important to do. Thus, this study aimed to describe the characteristics of Chemistry test items on USBN in Yogyakarta City.

METHOD

This research employs a descriptive exploratory study with quantitative and qualitative approaches. Quantitative data analysis is used to describe the characteristics of items using the classical test theory and item response theory. Qualitative data analysis was presented to describe the questions that belonged to difficult and not good categories. Respondents in this study were 194 students who took USBN for Chemistry subjects in Yogyakarta City.

Data were collected by documenting student responses or answers. The data source comes from the results of the implementation of USBN in Yogyakarta City which were consisted of 35 items presented in form of multiple-choice, students' answers, and question-answer keys. Quantitative data analysis uses Anbuso 4.0 to determine the level of difficulty and discrimination index of questions based on classical test theory. Meanwhile, Quest Program was used as supportive tool to find out the level of difficulty and reliability based on item response theory. The steps in qualitative analysis were identifying items that belonged to difficult category, outlining the problem-solving procedures, and analyzing the factors that might become the factor which causes difficulties and students' problems.

Data analysis with classical test theory was used to determine the level of difficulty and discrimination index of the questions, while item response theory was used to determine the level of difficulty and reliability value of USBN questions in Yogyakarta City. The level of difficulty and discrimination index of questions were identified based on the results of student answers in the implementation of USBN. Based on the analysis of classical test theory, questions were categorized as difficult if the coefficient of the difficulty index was below 0.3, easy questions with a coefficient of more than 0.7, and the rest were classified in the moderate category (Kolte, 2015, p.321). Discrimination index classified good and bad questions based on the results of students' answers with the coefficient of distinguishing value which was more than 0, and belonged to good category, while below 0 was categorized as bad category (Kolte, 2015, p.321). Based on the analysis of item response theory, the problem was categorized as difficult if the coefficient of difficulty index is more than 2, and the categorized as easy if it was less than -2, and category is in between (Sarea & Hadi, 2015, p.38).

FINDINGS AND DISCUSSION

Data analysis in this study used classical test theory analysis and item response theory. The result of reliability estimation based on item response theory was 0.48, which was categorized as less reliable category (Azwar, 2008, p.113). The results of the analysis of the level of difficulty and discrimination index based on the results of classical test theory by using Anbuso 4.0 are presented in Table 1.

Table 1. Difficulty Index Coefficient Based on Classical Test Theory

No	Difficulty Index Coefecient	Note	No	Difficulty Index Coefecient	Note	No	Difficulty Index Coefecient	Note
1	0.758	Easy	13	0.418	Moderate	25	0.361	Moderate
2	0.521	Moderate	14	0.232	Difficult	26	0.464	Moderate
3	0.448	Moderate	15	0.397	Moderate	27	0.057	Difficult
4	0.485	Moderate	16	0.727	Easy	28	0.582	Moderate
5	0.335	Moderate	17	0.201	Difficult	29	0.680	Moderate
6	0.897	Easy	18	0.624	Moderate	30	0.381	Moderate
7	0.789	Easy	19	0.840	Easy	31	0.809	Easy
8	0.773	Easy	20	0.314	Moderate	32	0.768	Easy
9	0.789	Easy	21	0.510	Moderate	33	0.541	Moderate
10	0.892	Easy	22	0.804	Easy	34	0.536	Moderate
11	0.768	Easy	23	0.510	Moderate	35	0.887	Easy
12	0.557	Moderate	24	0.289	Difficult			

Based on the difficulty index coefficient, it was found that the average level of difficulty of USBN questions is 0.57 which is included in the medium category based on classical test theory. This means that overall the Chemistry Subject USBN questions in Yogyakarta City have a medium category to be worked on by students. Judging from the magnitude of the difficulty index coefficient, the most difficult question order is item 27, 17, 14, and 24.

Table 2. Discrimination Index Coefficient Based on Classical Test Theory

Discrimination			Discrimination			Discrimination		
No	Index Coefficient	Note	No	Index Coefecient	Note	No	Index Coefficient	Note
1	0.027	Good	13	-0.086	Poor	25	0.146	good
2	0.051	Good	14	0.227	Good	26	0.110	good
3	0.141	Good	15	0.054	Good	27	-0.072	poor
4	0.167	Good	16	0.264	good	28	0.129	good
5	0.152	Good	17	0.100	good	29	0.273	good
6	0.065	Good	18	0.072	good	30	0.109	good
7	0.229	Good	19	0.323	good	31	0.247	good
8	0.094	Good	20	-0.041	poor	32	0.184	good
9	0.182	Good	21	0.283	good	33	0.234	good
10	0.146	Good	22	0.331	good	34	0.282	good
11	0.316	Good	23	0.016	good	35	0.161	good
12	-0.007	Poor	24	0.213	good			

Based on the coefficient of discrimination index in Table 2, the average discrimination index of the USBN problem was 0.146 which was included in both categories. This means that overall, the Chemistry test items in USBN Yogyakarta City belonged to good categories to be done by students, based on the coefficient of discrimination index, the order of the most unfavourable items is item 13, 27, 20, and 12. Based on the results of the coefficient of difficulty and the discrimination index of USBN questions, the data can be recapitulated as in Table 3.

Table 3. Data Recapitulation Based on Classical Test Theory Analysis

Difficulty Index Coefficient			Discrimination Index Coefficient	
Easy (11.43%)	Moderate (51.43%)	Difficult (37.14%)	Good (88.57%)	Poor (11.43%)
13 items	18 items	4 items	31 items	4 items
Number 1, 6, 7, 8, 9, 10, 11, 16, 19, 22, 31, 32, and 35	Number 2, 3, 4, 5, 12, 13, 15, 18, 20, 21, 23, 25, 26, 28, 29, 30, 33, and 34	Number 14, 17, 24, and 27	Number 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 28, 29, and 30	Number 12, 13, 20, and 27

Based on the results of the analysis with the help of Anbuso 4.0, four questions belonged to the difficult category. These problems consisted of material on salt hydrolysis, calculation of changes in reaction enthalpy, the process of making and use of chemical compounds, and the periodicity of the main metal elements.

Table 4. Difficulty Index Coefficient Based on Item Response Theory

No	Difficulty Index Coefecient	Note	No	Difficulty Index Coefecient	Note	No	Difficulty Index Coefecient	Note
1	-1.33	Moderate	13	0.2	Moderate	25	0.44	Moderate
2	-0.23	Moderate	14	1.1	Moderate	26	0.01	Moderate
3	0.07	Moderate	15	0.29	Moderate	27	2.74	Difficult
4	-0.08	Moderate	16	-1.16	Moderate	28	-0.52	Moderate
5	0.56	Moderate	17	1.28	Moderate	29	-0.95	Moderate
6	-2.36	Easy	18	-0.67	Moderate	30	0.36	Moderate
7	-1.51	Moderate	19	-1.85	Moderate	31	4.06	Difficult
8	-1.41	Moderate	20	0.67	Moderate	32	2.48	Difficult
9	-1.51	Moderate	21	-0.22	Moderate	33	-0.32	Moderate
10	-2.31	Easy	22	-1.6	Moderate	34	1.98	Moderate
11	-1.41	Moderate	23	-0.2	Moderate	35	2.95	Difficult
12	-0.38	Moderate	24	0.8	Moderate			

Problems included in the unfavourable category were four items with material on acid-base titration, the solution colligative phenomenon, and the concept of periodicity of the main metal elements. After analyzing the data in classical test theory, the results of the coefficient of difficulty according to item response theory are presented in Table 4.

Based on the difficulty index coefficient, it is found that the average level of difficulty of USBN questions is -0.00086 which is included in the moderate category based on item response theory. This means that overall the Chemistry Subject USBN questions in Yogyakarta City have a medium category to be worked on by students. Judging from the difficulty index coefficient, the most difficult question sequence is item 31, 35, 27, and 32. The recapitulation of the results of the difficulty index analysis based on item response theory is presented in Table 5.

Table 5. Recapitulation of Difficulty Index Based on Item Response Theory

Difficulty Index Coefcient		
Easy (5.71%)	Medium (82.86%)	Difficult (11.43%)
2 items	29 items	4 items
Number 6 and 10	Number 1, 2, 3, 4, 5, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 33, and 34	Number 27, 31, 32, and 35

Based on the results of the quantitative data analysis, there were eight difficult items and four bad items. One of the items in item 27 belongs to the three, which belonged to difficult category based on classical test theory and item response theory. After conducting quantitative data analysis, it was continued with a qualitative analysis of questions in order to describe the causes of the problems (the difficult and bad categories). Several examples of difficult and bad items based on classical test theory and item response theory are as follows.

Item 27

Item 27 was the most difficult item based on classical test theory analysis and the second hardest based on item response theory. This problem was also categorized as a bad question. The items presented are as follows.

Consider the physical and chemical properties of three elements such as the following unknown period:

Element	Boiling Point	Electrical Conductivity	Ionization Energy	Form
K	280°C	does not conduct electricity	1,012 kJ/mol	solid
L	2470°C	conduct electricity	0,579 kJ/mol	solid
M	58°C	does not conduct electricity	1,21 kJ/mol	gas

Arrange them based on the increasing of the atomic number of the element

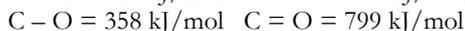
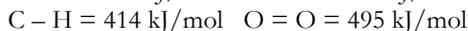
- a. M – L – K
- b. M – K – L
- c. K – M – L
- d. L – K – M
- e. L – M – K

The problem presented in the table of physical and chemical properties of an element in one period in the periodic table of elements, consisted of the initials of elements, boiling points, electrical conductivity, ionisation energies, and elemental states. Students were asked to arrange the three elements based on the increase in atomic number. The difficulty index coefficient for this problem was 0.057, in the difficult category. Interestingly, this problem was also included in the unfavourable category with a differentiating coefficient of -0.072. The negative value indicated that the problem was not good, since the table information was incomplete for students to answer the question. The tables containing boiling points and electrical conductivity should not be needed in determining the atomic number of an element. Periodicity consists of ionisation energy, atomic radius, and electronegativity which is an indicator to guess the atomic number of an element and its location, since the information in the table was insufficient to be inserted in a table of atomic radii and electronegativity which caused a difficult problem for answered and questions that were not good for students.

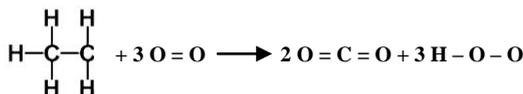
Item 17

After point 27, the next most difficult item was item 17 based on classical test theory. The items presented are as follows.

The data of average bond energy are:



Spiritus which contains ethanol is burned by the reaction equation:



The resulting enthalpy change for burning 23 grams of ethanol is

(Ar H = 1; C = 12; O = 16)

- a. -2500 kJ
- b. -1250 kJ
- c. -625 kJ
- d. +625 kJ
- e. +1250 kJ

This problem presented the equation of the reaction of spiritus containing ethanol. Students were asked to calculate the change in enthalpy of the combustion reaction with 23 grams of ethanol, while the average bond energy data had been known. The coefficient of the difficulty index of this problem was 0.201, categorized as difficult. This problem became difficult due to the problems which had relatively long calculation stages and required more accuracy, while the time given for each problem were relatively less.

Item 14

After items 27 and 17, the next most difficult item was item 14 based on the analysis of classical test theory. The items are presented as follows.

Consider the hydrolysis reaction data for the following salts:

No	Salt Formula	Type of Hydrolysis	Hydrolysis Reaction	Estimated pH
1	MgCl ₂	Total	$\text{MgCl}_{2(aq)} + 2\text{H}_2\text{O}_{(l)} \rightleftharpoons \text{Mg}(\text{OH})_2 + 2\text{H}^+_{(aq)} + 2 \text{Cl}^-_{(aq)}$	7
2	NaCH ₃ COO	Partly	$\text{NaCH}_3\text{COO}_{(aq)} + \text{H}_2\text{O}_{(l)} \rightleftharpoons \text{Na}^+_{(aq)} + \text{CH}_3\text{COOH}_{(aq)} + \text{OH}^-_{(aq)}$	>7
3	NH ₄ CH ₃ COO	Total	$\text{NH}_4\text{CH}_3\text{COO}_{(aq)} + \text{H}_2\text{O}_{(l)} \rightleftharpoons \text{NH}_4\text{OH}_{(aq)} + \text{CH}_3\text{COOH}_{(aq)}$	7
4	(NH ₄) ₂ SO ₄	Partly	$(\text{NH}_4)_2\text{SO}_{4(aq)} + 2\text{H}_2\text{O}_{(l)} \rightleftharpoons \text{NH}_4\text{OH}_{(aq)} + \text{H}_2\text{SO}_{4(aq)}$	<7
5	K ₂ CO ₃	not hydrolyzed	$\text{K}_2\text{CO}_{3(aq)} + 2\text{H}_2\text{O}_{(l)} \rightleftharpoons 2\text{K}^+_{(aq)} + \text{H}_2\text{CO}_{3(aq)} + 2\text{OH}^-_{(aq)}$	>7

The pair of data that are correlated correctly between the formula, the type of hydrolysis, and the hydrolysis reaction are..

- a. 1 and 2
- b. 1 and 3
- c. 2 and 3
- d. 3 and 4
- e. 3 and 5

This problem presented a data table of the hydrolysis reaction of several salts consisting of columns of salt formula, type of hydrolysis, hydrolysis reaction, and estimated pH. Students were asked to find the correct data pair for each column by selecting two pairs. The coefficient of the difficulty index of this problem is 0.232 where the questions belonged to the difficult category. This problem become difficult since it required a strong understanding of the concept of students to not be confused in interpreting the characteristics of salt, such as total hydrolysis, partial hydrolysis (partial), or not undergoing hydrolysis.

Item 24

Item 24 was the most difficult item because it has the lowest difficulty index compared to other items. The items presented are as follows.

The following table contains the formula for compounds, the manufacturing process, and uses of chemical compounds:

No	Compound Formula	Making Process	Usability
1	Na_2CO_3	Castner-Kellner method is by electrolyzing NaCl	Used as a water softener in washing clothes
2	NaOH	Solvay process	Used to remove rust stains on iron
3	NaClO	The Hooker process namely chlorine is passed cold and dilutes the hydroxide solution	Used in bleach liquid
4	NH_3	Harber process with a high temperature of $\pm 450^\circ\text{C}$ and pressure between 200-400 atm	Used as a raw material for the manufacture of rocket fuel hydrazine
5	H_2SO_4	Contact process with vanadium pentaoxide (V_2O_5) catalyst.	Used as a catalyst in the reaction of making alkenes

The correct pair of compound formulas, manufacturing processes, and uses of compounds are...

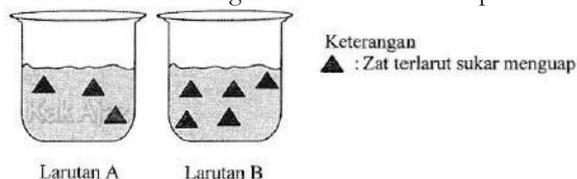
- a. (1), (2), and (3)
- b. (1), (3), and (4)
- c. (2), (3), and (5)
- d. (2), (4), and (5)
- e. (3), (4), and (5)

This problem presented a table containing the formulation of chemical compounds, the manufacturing process, and the uses of each chemical compound. Students are asked to match which chemical compounds are suitable for the manufacturing process and their use. Difficulty coefficient for this problem was 2,289 which belonged to the difficult category. This problem became difficult because it did not only require understanding but also required special memorization in the process of making compounds, while the names matched the names of the inventors.

Item 20

The next item that has a negative differentiation coefficient is item 20. Item 20 is presented as follows.

Consider the following illustration of the composition of the solution!



The most appropriate statement is...

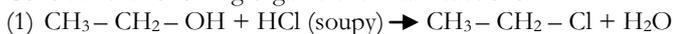
- a. Vapour pressure (A) is higher than the solution (B)
- b. The freezing point of solution (A) is higher than the solution (B)
- c. The boiling point of solution (A) is higher than the solution (B)
- d. The boiling point of solution (A) is the same as that of solution (B)
- e. The freezing point of solution (A) is the same as that of solution (B)

This problem presents an illustrated image of the composition of a solvent solution difficult to evaporate: solution A contains three substances, solution B contains five substances. Students were asked to identify which statement is the most appropriate among the five statements listed in the answer choice. Viewed from the difficulty index coefficient, this problem belongs to moderate category, but viewed from the differentiating power coefficient, this is considered as a bad question. The problem is similar to the problem that occurs in item 13. This is caused by to the concept of students still lacking for the material phenomenon of the colligative nature of the solution if presented with an image. Hence, even smart students can be distracted by the statement on the choice of answers to these questions, and the quality of the questions is not good.

Item 31

Item 31 was the item with the highest coefficient of blurring among the four other problems based on item response theory. Item 31 is presented as follows.

Consider the following organic chemical reactions!



The elimination reaction is in the number reaction equation...

a. (1) h. (4)

f. (2) i. (5)

g. (3)

The problem presented several reaction equations where students were asked to choose which reaction equation which showed an elimination reaction. Under the concept of the elimination reaction was a reaction that removed two substituents from a molecule. This reaction was usually characterized by the change of a single bond into a double bond by releasing small molecules. In the choice of answers, it appeared that the reaction equation (2) was showed. The possibility of this problem became difficult for students was the lack of understanding of the concept of organic chemistry.

Based on the results of quantitative analysis based on classical test theory and grain response theory, students' difficulties in chemistry subjects were found on the material nature of periodicity, salt hydrolysis, enthalpy of reaction, manufacture and use of chemical compounds, organic chemical reactions, formation of petroleum fractions, and structure, nomenclature, properties, and classification of macromolecules. If seen from the presentation of the items, students had difficulty in understanding the concept of the nature of periodicity, salt hydrolysis, and organic chemical reactions. Another difficulty was in calculating the change in reaction enthalpy, while in other materials, students had difficulty in answering the items because of the lack of students' insights, since the material required students' memorization abilities.

Based on the analysis results, many students had difficulty in answering questions related to chemical concepts and calculations. One of the chemical concepts that were difficult for students in Yogyakarta City to answer in the USBN was the chemical concept of salt hydrolysis material. Chemical equilibrium materials, acid-base solutions, and buffer solutions, which are the basic material for studying salt hydrolysis material. The lack of understanding of these material concepts makes students experience misconceptions on the material of salt hydrolysis. This is relevant to [Maratusholihah et al. \(2017, p.919\)](#) that chemical equilibrium and acid-base material are prerequisite materials for buffering solutions and salt hydrolysis. Other studies also revealed that students will experience difficulties in understanding chemical materials that require understanding the concepts of chemical equilibrium prerequisites ([Indriani et al., 2017, p.10](#)).

The misconception was caused by material, teacher, or student factors. Material that tends to be abstract and complex makes it difficult for students to understand salt hydrolysis material. In addition, chemistry learning which is often more oriented to chemical calculations to be able to answer the problem well finally put aside the basic concepts that should be understood first before answering the problem. This is relevant to research conducted by [Maratusholihah et al. \(2017, p.919\)](#), which revealed that chemical materials require integration in the macroscopic, microscopic, and symbolic aspects of studying buffer material and salt hydrolysis. Errors caused by the lack of integration of these aspects lead to misconceptions among students that are caused by students' daily experiences, textbooks for subjects, and learning by the teacher.

One solution that can reduce misconceptions among students is to improve the teaching method to provide more enjoyable learning so they do not get bored in the learning process. This had been studied by other researcher ([Akbar, 2016, p.6](#)) who was looking for solutions in solving the difficulties of students studying salt hydrolysis material based on the teacher's reflection. The results show that the teachers suggested extending the material to the modern applicative level,

increasing the practicum, procuring pretest with a maximum of three items, and using assessment tools using multiple-choice questions, and the construction of questions related to material outside hydrolysis. The material can use prerequisites for salt hydrolysis, namely, chemical equilibrium, acid-base and buffer solution. Based on the results of these studies indicate that the teacher has very large role in solving misconceptions in students, especially in the concept of chemistry.

Difficulties of other chemical concepts also occur in the material nature of periodicity. Students have difficulty in determining and sorting atomic numbers based on the nature of the periodicity listed. The nature of periodicity is a basic concept that must be understood when students study chemistry. This problem is very rarely found and researched by many researchers because the material is not categorized in difficult material and complained by students. Therefore, teachers and education experts can examine the problem as a new reference for teachers in teaching.

In addition, the concept of chemistry and chemical calculations were also the cause of the difficulty of the USBN problem in Yogyakarta City. The chemical calculations presented in the problem were the elliptical calculations of the reactions in the thermochemical chapter. This material required a deep understanding of the concept and calculation of chemical reactions. Thermochemistry is one of the chemicals that was considered as relatively difficult and less attractive to students because many use calculations. Errors often occur in calculations caused by poor understanding of concepts. The concept of thermochemistry was very closely related to the previous material, such as the concept of chemical equilibrium. The lack of the concept of chemical equilibrium in students will be an additional obstacle for studying thermochemical material. This was in line with research conducted by [Sugiawati \(2013, p.27\)](#) which found misconceptions on thermochemical material that made it difficult for students to identify a reaction equation and caused errors in the calculation of H reactions in the questions. This happened since one of them is due to the way the teacher submits that is only focused directly on chemical calculations and does not care much about planting students' concepts.

Many studies have examined these problems by holding innovations in overcoming the difficulties of thermochemical learning. One study was conducted by [Dirgahayuning \(2017, p.14\)](#) who implemented an active learning strategy with Learning Start With Question (LSQ) to achieve student learning completeness. This strategy seeks students to actively ask questions in the learning process, so students are required to be active in making questions before being explained by the teacher by emphasizing reading and asking skills before entering class.

In addition to the concept of hydrolysis, the nature of the periodicity and enthalpy of the reaction, some chemicals require memorization and the daily experience of students. As with the material process of making chemical compounds. After students understand the basic concepts of elements and compounds, students also need to know how the process of making these compounds. Difficulties also occur in the matter of organic chemical reactions, the formation of petroleum fractions, and the structure, nomenclature, properties, and classification of macromolecules. This becomes difficult because not all students can memorize well and lack students' insights. The problem is rarely found and not studied deeply by researchers, so teachers and education experts can examine the problem as a new reference in learning chemistry.

Problems in chemistry such as the lack of understanding of concepts, accuracy in counting, and memorization of certain materials. In addition to improving the learning system by designing better learning strategies, teachers can also provide more motivation for students to be more active in learning. Learning in class is sometimes still focused on the teacher, so some students feel bored, sleepy and indifferent. According to [Stuckey et al. \(2013, p.27\)](#), the teacher must realize that not all students have the same interests and pay attention to the consequences. Many students are still not intrinsically motivated by science learning. These conditions cause students' learning motivation to be low and will have an impact on students' cognitive achievement ([Lubis & Ikhsan, 2015, p.192](#)). Student learning motivation will increase with more structured learning. One way to increase motivation and learning outcomes of students is by giving structured assignments accompanied by providing feedback on direct learning ([Sabriani, 2012, p.41](#)). These struc-

tured tasks can help students and make them easy learning since they have been given clear learning syntax. In addition to structured tasks, giving feedback such as awards can also increase students' learning motivation.

Another factor that can support the process of learning chemistry is by instilling the significance of science learning to students. According to [Fitriani et al. \(2014, p.2\)](#), meaningfulness in learning can be grown from good chemical literacy skills also in students. It has also been studied which combines social-scientific problems, Life Cycle Analysis (LCA), and Inquiry-Based Learning (IBL) on chemistry teaching. The research has the opportunity to grow the goals of modern education including scientific literacy ([Juntunen & Aksela, 2013, p.150](#)). Good chemical literacy skills can be improved by connecting chemical concepts in everyday life, so learning becomes more meaningful and students can save long-term memory not just to answer questions.

Reviewing these problems not only lies in students but can be caused by teacher factors. This was shown based on the results of the analysis using the help of Quest, the reliability of the USBN questions in Yogyakarta City was 0.48, categorized as less reliable. According to [Azwar \(2008, p.113\)](#), the coefficient of reliability was in the range of 0.00 to 1.00. If the coefficient of reliability approaching 1.00 can be said to be more reliable, whereas if getting closer to 0.00 it is said to be less reliable. This makes the assessment standards less appropriate for measuring the ability of students in chemistry subjects, especially for questions of a national standard such as the USBN problem, because psychometric evidence is very important in making instrument ([Arjoon et al., 2013, p.536](#)). These problems can be solved by increasing teacher insight in preparing and conducting assessments, such as conducting training for teachers. The teacher can also conduct a study in the classroom to find solutions to improve effective and efficient learning that can improve students' motivation and learning outcomes.

CONCLUSION

The results show that the USBN device chemistry subject in terms of difficulty based on the classical test theory consisted of 13 easy items, 18 moderate items, and four difficult items and discrimination index consisted of 31 good items and four bad items. If it was reviewed based on the response theory, the item difficulty index consists of two easy items, 29 moderate items, and four difficult items. The estimated reliability results of 0.48 which indicates that the questions are less reliable. Problems with difficult categories consist of material properties of periodicity, salt hydrolysis, reaction enthalpy, manufacture and use of chemical compounds, organic chemical reactions, the formation of petroleum fractions, and structure, nomenclature, properties, and classification of macromolecules. Meanwhile, the problem of bad category consists of acid-base material, acid-base titration, colligative nature of the solution, and the concept of periodicity.

These materials were categorized to be difficult and not good because of the lack of understanding of students' concepts, lack of accuracy in counting, and difficult to memorize for certain chemical materials. Problems with the category of not good questions on this USBN problem were based on in the type of questions by choosing the right answer from a given phenomenon. These problems can be solved by improving appropriate learning strategies, providing structured assignments, and providing feedback to students, to make the learning process more meaningful in students' daily lives and can increase motivation and students' learning achievements.

The first thing in solving the easy, bad, and unreliable items is to review the teacher's ability, prepare, and conduct assessments. These results can be used as consideration for improvement, for example by conducting special training for teachers, especially in national assessments for students' graduation standards. It also can invite teachers and education experts to conduct research on learning chemistry that is more effective and efficient in the classroom, especially in creating material process and the use of chemical compounds, the concept of periodicity, and organic chemical reactions which are still considered difficult for students and are still very rare to be studied. This innovation is needed to improve the quality of education, especially in Chemistry subject, in Yogyakarta City.

ACKNOWLEDGMENT

The authors thank the Center of Assessment and Instruction, Indonesia, for supporting this research in terms of funding research and data retrieval. The research was also supported by the Chemistry teacher in Yogyakarta City regarding the USBN result data.

REFERENCES

- Akbar, S. A. (2016). Desain didaktis pembelajaran hidrolisis didasarkan hasil refleksi diri guru melalui Lesson Analysis. *Jurnal Edukasi Kimia*, 1(1), 6–11. <http://ojs.serambimekkah.ac.id/JEK/article/view/161>
- Ariani, L. H. D., Maimunah, M., & Roza, Y. (2019). Analisis kesalahan siswa dalam menyelesaikan soal USBN matematika SMA. *EduMath*, 8(1), 42–48. <https://ejournal.stkipjb.ac.id/index.php/math/article/view/1117>
- Arjoon, J. A., Xu, X., & Lewis, J. E. (2013). Understanding the state of the art for measurement in chemistry education research: Examining the psychometric evidence. *Journal of Chemical Education*, 90(5), 536–545. <https://doi.org/10.1021/ed3002013>
- Azwar, S. (2008). *Reliabilitas dan validitas*. Pustaka Pelajar.
- Dirgahayuning, A. (2017). Penerapan strategi pembelajaran aktif learning start with question untuk mencapai ketuntasan belajar siswa pada pokok bahasan termokimia kelas XI IPA 6 SMA Negeri 5 Pekanbaru. *Perspektif Pendidikan Dan Keguruan*, 8(2), 13–20. Retrieved from <https://journal.uir.ac.id/index.php/Perspektif/article/view/775>
- Fitriani, W., Hairida, H., & Lestari, I. (2014). Deskripsi literasi sains siswa dalam model inkuiri pada materi laju reaksi di SMAN 9 Pontianak. *Jurnal Pendidikan Dan Pembelajaran*, 3(1), 1-13. <https://jurnal.untan.ac.id/index.php/jpdpb/article/view/4432>
- Herkusumo, A. P. (2011). Penyetaraan (equating) ujian akhir sekolah berstandar nasional (UASBN) dengan teori tes klasik. *Jurnal Pendidikan Dan Kebudayaan*, 17(4), 455-471. <https://doi.org/10.24832/jpnk.v17i4.41>
- Indriani, A., Suryadharma, I. B., & Yahmin, Y. (2017). Identifikasi kesulitan peserta didik dalam memahami kesetimbangan kimia. *J-PEK (Jurnal Pembelajaran Kimia)*, 2(1), 9–13. <https://doi.org/10.17977/um026v2i12017p009>
- Juntunen, M., & Aksela, M. (2013). Life-cycle analysis and inquiry-based learning in chemistry teaching. *Science Education International*, 24(2), 150–166.
- Kolte, V. (2015). Item analysis of multiple choice questions in Physiology examination. *Indian Journal of Basic and Applied Medical Research*, 4(4), 320–326. <https://www.ijbamr.com/assets/images/issues/pdf/september%202015%20320-326.pdf.pdf>
- Lubis, I. R., & Ikhsan, J. (2015). Pengembangan media pembelajaran kimia berbasis Android untuk meningkatkan motivasi belajar dan prestasi kognitif peserta didik SMA. *Jurnal Inovasi Pendidikan IPA*, 1(2), 191-201. <https://doi.org/10.21831/jipi.v1i2.7504>
- Maratusholihah, N. F., Rahayu, S., & Fajaroh, F. (2017). Hidrolisis garam dan larutan penyangga. *Jurnal Pendidikan: Teori, Penelitian Dan Pengembangan*, 2(7), 919–926. <http://journal.um.ac.id/index.php/jptpp/article/view/9645/4559>
- Mardapi, D. (2016). *Pengukuran, penilaian, dan evaluasi pendidikan*. Parama Publishing.
- Retnawati, H. (2015). Hambatan guru matematika sekolah menengah pertama dalam menerapkan kurikulum baru. *Jurnal Cakrawala Pendidikan*, 34(3), 390–403. <https://doi.org/10.21831/cp.v3i3.7694>

- Ristiyani, E., & Bahriah, E. S. (2016). Analisis kesulitan belajar kimia siswa di SMAN X Kota Tangerang Selatan. *Jurnal Penelitian Dan Pembelajaran IPA*, 2(1), 18–29. <https://doi.org/10.30870/jppi.v2i1.431>
- Sarea, M. S., & Hadi, S. (2015). Analisis kualitas soal ujian akhir semester mata pelajaran Kimia SMA di Kabupaten Gowa. *Jurnal Evaluasi Pendidikan*, 3(1), 35–43. <http://journal.student.uny.ac.id/ojs/index.php/jep/article/view/1223>
- Sabriani, S. (2012). Penerapan pemberian tugas terstruktur disertai umpan balik pada pembelajaran langsung untuk meningkatkan motivasi dan hasil belajar siswa (Studi pada materi pokok struktur atom kelas X6 SMA Negeri Watampone). *Chemica: Jurnal Ilmiah Kimia dan Pendidikan Kimia*, 13(2), 39–46. <https://ojs.unm.ac.id/index.php/chemica/article/view/625>
- Stuckey, M., Hofstein, A., Mamlok-naaman, R., & Eilks, I. (2013). The meaning of ‘relevance’ in science education and its implications for the science curriculum. *Studies in Science Education*, 49(1), 1–34. <https://doi.org/10.1080/03057267.2013.802463>
- Sugiawati, V. A. (2013). Penggunaan strategi konflik kognitif dalam pembelajaran TPS untuk mereduksi miskonsepsi siswa pada materi termokimia. *Jurnal Nalar Pendidikan*, 1(1), 26–31. <https://ojs.unm.ac.id/nalar/article/view/1935>
- Taber, K. S. (2013). Revisiting the chemistry triplet: Drawing upon the nature of chemical knowledge and the psychology of learning to inform chemistry education. *Chemistry Education Research and Practice*, 14(2), 156–168. <https://doi.org/10.1039/c3rp00012e>
- Yustika, A., Susatyo, E. B., & Nuswowati, M. (2015). Uji kriteria instrumen penilaian hasil belajar kimia. *Jurnal Inovasi Pendidikan Kimia*, 8(2), 1330–1339. <https://journal.unnes.ac.id/nju/index.php/JIPK/article/view/4438>