Development of Knowledge Test Instruments of Prospective Bachelor of Biological Education Based on TPACK and KKNI Framework

Rabia Hidayanti Arnan, Evi Suryawati, Imam Mahadi
Magister of Biologi Education, University of Riau Pekanbaru, 28293, Indonesia

ARTICLE INFO

A B S T R A C T

This research is a descriptive study that aims to produce a test instrument for the knowledge of candidates for Biology Education undergraduate candidates for the Technologhy Pedagogy and Content Knowledge (TPACK) and Indonesian National Qualifications Framework. The study was a development with plomp models conducted at the stage of development. Data collection was done by validation and testing. Validation aims to look at the alignment of experts' opinions, while testing serves to see the handling of the problem by analyzing the test results. Data obtained from test results analysis using rasch modelling with the help of the winstep software. The results of the analysis are in the form of the value of reliability, unidimensionality, level of difficulty, distinguishing power, distractor and person - item map. The results showed that the number of valid questions was 63 items, the number of questions that were invalid or not accepted was 17 items. Invalid questions must be followed up in order to produce a good knowledge test instrument so that it can be used in measuring the mastery of TPACK for prospective teachers and biology teachers.

1. Introduction

The 21st century is marked by changes in various aspects of human life, one of which is marked by the use of information and communication technology, including in the learning process. The world of work demands a change in competence. An important competency in the 21st century is characterized by the ability to think critically, solve problems and collaborate. Educators and Education Personnel Institutions (LPTK) are the highest education institutions that carry out undergraduate and postgraduate programs as well as professions that are directed primarily at mastering educational disciplines, in accordance with

the Indonesian National Qualifications Framework (KKNI) which is able to prepare students as prospective teachers who can meet the needs of current education personnel.

This is certainly a tough challenge for the Indonesian education world. Moreover, in the global context also developed a form of teacher competency development pattern called "TPACK" or Technological, Pedagogical, Content Knowledge which is the point, a teacher must have comprehensive and holistic knowledge and skills in terms of content/materials, pedagogy/educational science and technology (Nofrion et al, 2018). Desmita et al (2021) also said that the level of motivation and learning achievement of students is also related to the pedagogical competence and professional competence of a teacher. Either the bad quality of the pedagogical competence of teachers will reduce or increase student learning motivation and also impact student learning discipline (Marina et al, 2019). Mastery of teacher TPACK skills greatly affects the quality of learning in the class. This is because TPACK's ability to influence teacher confidence in teaching (Arsal, 2014). The ability of TPACK teachers needs to be constantly developed and even adapted to the technological era so that it will be able to increase the level of self-efficacy.

Baser et al (2015) says that instrument development is a good way to see a person's TPACK more specifically, although there are some difficulties in making instrument items. TPACK competency test for prospective biology teachers conducted on 4 LPTK in Riau Province in 2015 got low results with an average overall success percentage of only 40.8% (Suryawati and Roza, 2017). This is due to the lack of preparation of participants in preparing to take the exam because the participants are not trained in working on questions or discussing questions. Yeh et al (2014) has also conducted research aimed at developing technology literacy measuring instruments, but the item of instrument items he developed is still not perfect because it has not been able to demonstrate the ability of teachers in organizing pedagogical abilities towards the literacy of technology that corresponds to the topic of learning materials. Yulvisriani et al (2020) research suggests that there are still a few teachers who do not show competence in teacher competence, as it requires a promotion to increase a teacher's confidence in improving each teacher's competence.

LPTK as an academic education organizer and profession needs to conduct studies related to improving the quality of education. The assessment of instruments that can be used in measuring TPACK can help in improving knowledge competencies for prospective teachers. The development of knowledge test instruments based on the framework of TPACK and KKNI will be able to be used by prospective bachelors of biology education and for those who want to continue to the teacher profession education program to train and improve the ability to work on exam questions. Therefore, this research aims to produce The Knowledge Test Instrument of Prospective Bachelor of Biological Education Based on technological pedagogical and content knowledge framework (TPACK) and KKNI.

2. Methodology

This study used a development research for prospective undergraduate students of Biology Education to produce a knowledge test instrument based on TPACK and KKNI. The development model used in this study is plomp model with research stage consisting of top, preliminary research phase, development or prototyping phase and assessment phase, where research is only done until the stage of prototype creation (development). This research was conducted at FKIP Biology University of Riau with a sample 30 prospective undergraduate students of biology education.

The instruments used in this study are validation sheet instruments and designed instruments. The validation sheet aims to see the usefulness of the questions from the suggestions and input of experts in the field of pedagogics and biological content. The instrument designed is 80 items based on the framework of TPACK and KKNI. Technically assisted data analysis using Winstep software. The results of the analysis are in the form of the value of reliability, unidimensionality, level of difficulty, distinguishing power, trick function and person - item map.

3. Results and Discussion

This research begins by doing analysis. At the analysis stage researchers do some need analysis. The analysis is analysis of the study of theory / relevant research and analysis of the curriculum of the faculty of biology education FKIP UNRI. This analysis is the basis in formulating grid indicators from the development of knowledge test instruments of prospective biology education scholars based on the framework of TPACK and KKNI. The results of the analysis in the form of grid design and mapping of instrument indicators test the knowledge of prospective biology education scholars based on the framework of TPACK and KKNI.

The problem is in the field of Biology Education, namely Pedagogy and Biological Content. The cognitive level used is adjusted to Bloom taxonomy with a ratio of 25% easy question: 50% medium : 25% difficult. After the analysis phase, the next stage of development is carried out. At this stage, researchers began designing knowledge test instruments for prospective biology education scholars based on grids that have been made in accordance with the frameworks of TPACK and KKNI.

The instrument developed in the form of cognitive questions is a multiple choice with 5 answer options or 4 distractors totaling 80 questions. After the design of the question at the development stage, then the next stage is the researcher develops the design into a product that is a test instrument of knowledge prospective bachelor of biology education.

Validation Results of Knowledge Test Instruments of Prospective Biological Education Scholars Based on TPACK and KKNI Framework

Table 1 above shows the recapitulation of the average gain of validity scores for all aspects of validity. The average material aspect score is 3.75 with a very valid category, for construction aspect is 3.79 with very valid category, and for language aspect of 3.75 with very valid category, so the average score for all aspects of validity is 3.76 with very valid's category. The score shows that the instrument designed as a whole is valid, making it eligible to be used and applied to prospective biology education scholars. According to Majid (2014), tests with high validity mean being able to express certain aspects of learning appropriately. Instruments that have met the validity criteria are further tested.

Table 1. Average recapitulation of validity score of knowledge test instrument prospective biology education undergraduate based on TPACK and KKNI framework

No	Assessment Aspect	Average Score				Average Score	Validity Category
		V – 1	V – 2	V – 3	V – 4		
1	Material	3.45	3.82	3.73	4.00	3.75	Very Valid
2	Construction	3.71	3.86	3.57	4.00	3.79	VeryValid
3	Language	3.75	3.75	3.50	4.00	3.75	Very Valid
Average Total Score						3.76	Very Valid

Description:
V-1: Validator 1; V-2: Validator 2; V-3: Validator 3; V-4: Validator 4

Results of Development of Knowledge Test Instruments of Prospective Bachelor of Biological Education Based on TPACK and KKNI Framework

A. Reliability

Reliability of test instruments developed in this study based on alpha cronbach is with a value of 0.97 which means the reliability score between students and the question item belongs to a very good category (Table 2). Reliability of the question items that have reached a modest high indicates that this instrument is adequate enough and can be used to conduct real research (Hayati & Lailatussaadah, 2016).

Table 2. Instrument Reliability Test Results

Variabel	Mean Measure	Separation	Reliability	Alpha Cronbach
Students	-0,40	5,04	0,96	0,97
Items	0,00	2,78	0,89	

The separation value obtained can also be demonstrated as the quality of the instrument as well as the quality of the subjects of this study. The student separation score is 5.04 and the separation point question value is 2.78. The greater the separation value, the more able to show the quality of the instrument grain question used very well (Sumintono & Widhiarso, 2014).

This is because it is able to identify the group of respondents (able-able) and the broader group of (difficult-easy) questions. The formula commonly used to view

groupings more thoroughly is called strata separation with formula $H_{Students}=$ [(4*separation)+1]/3 then $H_{Students} =$[(4*5,04)+1]/3, $H_{Students} = 7.05$ or rounded to 7. $H_{items}=$[(4*separation)+1]/3 then $H_{items}=$[(4*2,78)+1]/3, $H_{items}= 4.04$ or rounded into 4, which separation question is good value because it can divide the question into 4 groups. Based on the guidelines for instrument quality developed by Fisher, W.P Jr (2007), it can be seen in the image below that this group of developed question items can be classified in the very good category because it reaches a value of 4. Lidinillah et al (2020) in his research also said about the value of separation, which is known as the criterion that the greater the value of separasi then the better the quality of the instrument.

B. Unidimensionality

The results of analysis of the knowledge test instrument obtained raw variance measurement value of 53.6% and fall into the category of good. This indicates that a minimum unidimensionality requirement of 20% can be met (Sumintono & Widhiarso, 2014). Thus the multiple choice test instruments developed in this study are valid for measuring the abilities of students.

C. Difficulty level

Table 3 shows the results of the analysis of the difficulty level of each item of tpack knowledge test instrument, shows the level of difficulty for each category does not match the one stipulated. Question items that fall into the category are easily obtained 15.00% percentage (12 question items), moderate category with a percentage of 48.75% (39 question items), and difficult category with a percentage of 36.25% (29 question items). This shows that the distribution of questions is uneven or unbalanced because the proportion of difficulty level is balanced with the curve of 25% (easy): 50% (medium): 25% (hard). This means that more difficult items for students were as much as 36.25%. Meanwhile, the easy items for students were only 15.00%, meaning that some of the questions that were categorized as easy were included in the medium and hard category for students who worked on them.

Table 3. Proportion and Percentage of Difficulty Level Of Grains About Knowledge Test Instruments prospective Bachelor of Biology Education

No	Category	Competence	Amount	Percentage
1	Easy	*Technological Knowledge* (TK)	0	15,00%
		Pedagogical Knowledge (PK)	6	
		Content Knowledge (CK)	1	
		Technological Pedagogical Knowledge (TPK)	2	
		Technological Content Knowledge (TCK)	1	
		Pedagogical Content Knowledge (PCK)	2	
		Technology Pedagogical and Content Knowledge (TPACK)	0	
2	Medium	*Technological Knowledge* (TK)	1	48,75%
		Pedagogical Knowledge (PK)	3	

Content Knowledge (CK)	31	
Technological Pedagogical Knowledge (TPK)	0	
Technological Content Knowledge (TCK)	1	
Pedagogical Content Knowledge (PCK)	3	
Technology Pedagogical and Content Knowledge (TPACK)	0	
Technological Knowledge (TK)	0	
Pedagogical Knowledge (PK)	3	
Content Knowledge (CK)	22	
Technological Pedagogical Knowledge (TPK)	1	36,25%
Technological Content Knowledge (TCK)	1	
Pedagogical Content Knowledge (PCK)	1	
Technology Pedagogical and Content Knowledge (TPACK)	1	

(Row number 3 appears at left for the second group)

According to Muslich (2011) a good test instrument is an instrument that has an even level of difficulty. Empirically the question with low and high difficulty levels is less able to distinguish the position of participants with the category of clever and weak. Therefore, it is recommended that the question that is widely used is a question with a medium and upper difficulty level.

D. Differentiating power of items

In Table 4 percentage of differentiating power for each question item, the good category has a percentage of 13.75% (11 items), a good category of 35.00% (28 items), a category of just 32.50% (26 items). For this reason, this question item can be accepted because it can distinguish high, low and medium ability test takers. Items in the bad category were 13.75% (11 items) and only 5% negative categories (4 items). This shows that the problem cannot yet be categorized as an acceptable question because it has not been able to distinguish between upper-class test takers and lower-class test takers.

Table 4. Proportion and Percentage of Grain Differentiation Power of Knowledge Test Instruments prospective Bachelor of Biology Education Test Stage

No	Category	Problem Deskrimination Index	Competence	Amount	Percentage
1.	Very Good	0,70 – 1,00	*Technology*	0	13,75%
			Pedagogic	1	
			Content	9	
			TPK	0	
			TCK	1	
			PCK	0	
			TPACK	0	
2.	Good	0,40 – 0,69	*Technology*	0	35,00%
			Pedagogic	6	
			Content	21	
			TPK	1	

				TCK	0	
				PCK	0	
				TPACK	0	
3.	Sufficient	0,20 – 0,39		*Technology*	0	32,50%
				Pedagogic	3	
				Content	17	
				TPK	1	
				TCK	2	
				PCK	3	
				TPACK	0	
4.	Bad	0.00 – 0,19		*Technology*	1	13,75%
				Pedagogic	1	
				Content	5	
				TPK	1	
				TCK	0	
				PCK	1	
				TPACK	1	
5.	Discarded	Negative (-)		*Technology*	0	5,00%
				Pedagogic	1	
				Content	2	
				TPK	0	
				TCK	1	
				PCK	1	
				TPACK	0	

Items with a negative discriminatory power index value are then replaced by new items. Daryanto (2005) stated that the power of different questions is the ability of a question to distinguish highly capable test taker with low ability. In line with the statement, Arikunto (2015) stated that the question that is well received is a question that has different power categories accepted. Therefore, it is necessary to follow up on the question.

E. Distractor

The results of the analysis of the function of the beggar on the knowledge test instruments of prospective biological education scholars based on the framework of TPACK and KKNI shows that the problem already has a good and effective problem-testing function but there are still some problem-ers who do not work or do not work. Some of the questions that do not work well include numbers 23, 29, 44, 39, 64, 20, 16, 4, 27, 11, 19, 63.

Table 5. Analisys Distractor

```
-------------------------------------------------------------------------------
|ENTRY   DATA  SCORE |    DATA   |     ABILITY      S.E.  INFT OUTF PTMA |      |
|NUMBER  CODE  VALUE |  COUNT  % |   MEAN    P.SD  MEAN   MNSQ MNSQ CORR.| Item |
|--------------------+-----------+---------------------------------------+------|
|  50    E        0  |    9  30  |   -.75    1.33  .47    .3   .5  -.12 |S50   |
|        C        0  |   10  33  |   -.44    2.96  .99   2.0  2.7  -.02 |      |
|        A        0  |    5  17  |    .34     .15  .07    .3   .7   .17 |      |
|        D        0  |    3  10  |    .43     .07  .05    .3   .8   .14 |      |
|        B        1  |    3  10  |  -1.26*    .38  .27  10.0  7.4  -.15 |      |
|                    |           |                                       |      |
|  51    E        0  |    4  13  |  -3.92     .22  .13    .0   .0  -.71 |S51   |
|        B        0  |    2   7  |  -1.00     .13  .13    .1   .2  -.08 |      |
|        C        0  |   11  37  |    .20     .64  .20    .3   .7   .23 |      |
-------------------------------------------------------------------------------
```
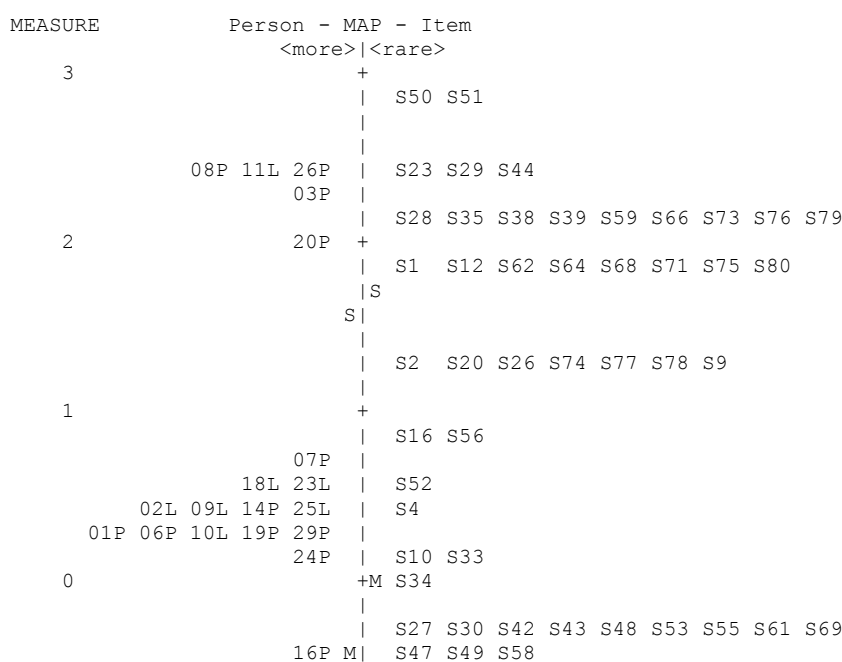
```
|      A        0 |    10  33 |    .23    2.18  .73  1.9  2.8    .23 |       |
|      D        1 |     3  10 |    .43     .07  .05  3.5  1.3    .14 |       |
|                |           |                                     |       |
-----------------------------------------------------------------------------
```

On item no 50, the correct option is B. Option E answered is 9 people, option C there are 10 people, option A there are 5 people and option D there are 3 people and who answer option B there are 3 people who answer correctly. This analysis of the fool works judging by its Mean value. If the Mean value rises then the picker or distractor works fine. At question no 50, the mean value does not rise then the distractor does not work properly. This means that some low-skilled people can answer questions correctly so that the distractor doesn't work.

The problem with functioning speed is a good question because the function of the tester on the question can trick test takers who do not understand the material being tested. Sulistiawan (2016) stated that the effective function of coding is caused by the preparation of sentences in the choice of answers, making it easier for test takers who understand the material to answer questions appropriately.

F. Person – item map

Based on the results of the analysis using winsteps program provides information, both in terms of items and respondents show differences in question items and students analyzed using rasch model. According to Arikunto (2013) stated that a question that can be answered correctly by clever students and less good students is a problem that is not good because it has no distinguishing power. Figure 1 person-item map shows the spread of the ability of 30 students (left) and the spread of difficulty level of question grains (right) on the same scale. The results of the data analysis in the image above are obtained information on the extent to which the details of the test constituents developed are worth using to measure the ability of students.

```
MEASURE           Person - MAP - Item
                    <more>|<rare>
  3                     +
                        |  S50 S51
                        |
                        |
          08P 11L 26P  |  S23 S29 S44
                  03P  |
                        |  S28 S35 S38 S39 S59 S66 S73 S76 S79
  2                20P  +
                        |  S1  S12 S62 S64 S68 S71 S75 S80
                       |S
                     S|
                        |
                        |  S2  S20 S26 S74 S77 S78 S9
                        |
  1                     +
                        |  S16 S56
                  07P  |
             18L 23L  |  S52
         02L 09L 14P 25L  |  S4
    01P 06P 10L 19P 29P  |
                  24P  |  S10 S33
  0                   +M S34
                        |
                        |  S27 S30 S42 S43 S48 S53 S55 S61 S69
                 16P M|  S47 S49 S58
```

```
                         |
                         |  S11 S8
                 15P     |  S19 S36 S41 S63
   -1                    +
                 30L     |  S25 S31 S32 S37 S46 S57 S60 S65 S67 S70 S72
                         |
                         |  S3  S40 S54
                         |
                 17P     |S S18 S45
                         |
   -2                    +
          04P 12P 21P    |
                 27P S|  S13 S5   S6   S7
                         |
                         |
                         |
                         |
   -3                    +  S24
                         |
                         |
                         |T S14 S15 S17 S21 S22
                 05P     |
                         |
                 22P     |
   -4                    +
             13P 28P     |
                     T|
                         |
                         |
                         |
                         |
   -5                    +
                 <less>|<freq>
```
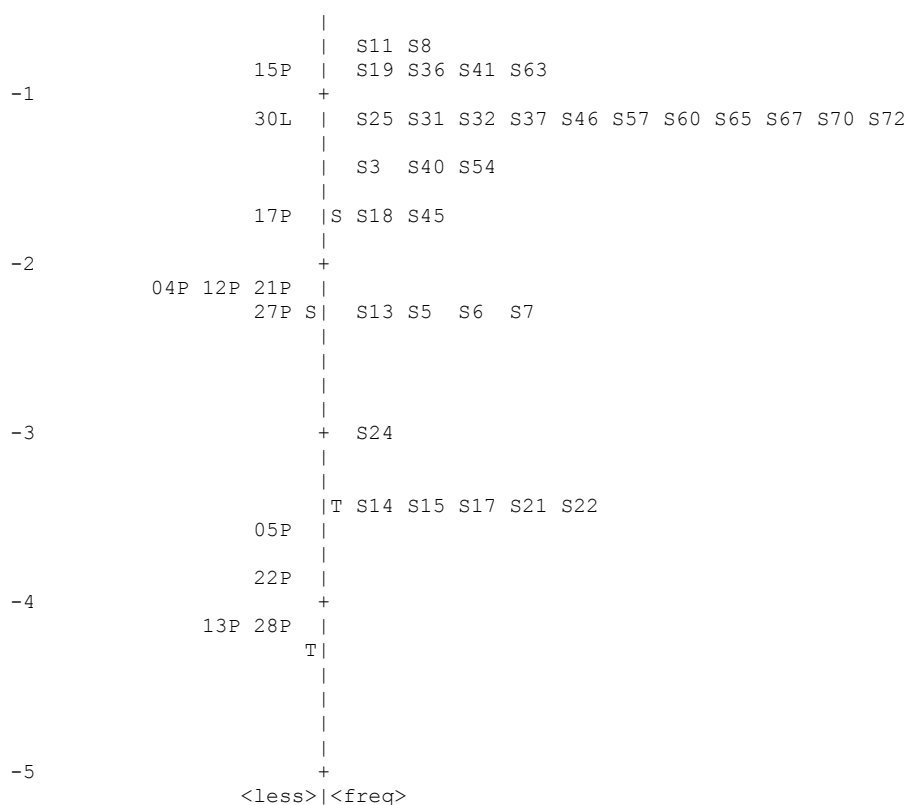
Figure 1. *Person–Item Map*

The results of the data analysis in picture 2 shows that there are three highly capable students, namely students with codes 03P, 08P, 11L, 20P, and 26P with a logit value achieved +2, this indicates that the student already understands and answers the question that has been tested well, from the data can also be known that some students are in intermediate ability with a logit value of +1 , and there are still many students with low skills with a logit value of -2 down which means there are still some students who have not been able to answer the question correctly. Students with 05P, 22P, 13P, and 28P codes are at the very bottom, even if the easy questions are not biased to answer correctly. This can be due to the lack of seriousness of students in working on the questions that have been given.

The results of the analysis on the knowledge test instrument of prospective biology education scholars based on the framework of TPACK and KKNI obtained that the knowledge test instrument developed is good enough but there are some details of the question that are not yet valid. The number of valid questions is 63 questions while the number of invalid or unacceptable questions there are 17 questions found on the question number (1, 11, 12, 16, 19, 20, 23, 27, 29, 39, 44, 50, 51, 63, 64, 77, 80). The invalid problem continued with a follow-up fixed to the perfection of the instrument of the question of tinging. Valid questions can be stored as a question/test bank and can be used to measure the mastery of TPACK of prospective teachers and biology teachers.

4. Conclusion

Based on the research that has been done, it can be concluded that the instrument developed is already in a very valid category with a value of 3.76. Reliability score of 0.97 with very good category and unidimensionality value of 53.6% with good category. But there has to be an improvement in the problem because the level of difficulty in the problem still does not have an even spread. The differentiating power of the developed problem also still exists negative value. Therefore it is necessary to follow up on the developed question.

Acknowledgments

References

Arikunto, S. (2015). *Dasar-Dasar Evaluasi Pendidikan Edisi 2*. Jakarta: Bumi Aksara

Arsal, Zeki. (2014). Microteaching and pre-service teachers' sense of self-efficacyin teaching. *European Journal of Teacher Education*, *37(4),* 453–464. http://doi.org/10.1080/02619768.2014.912627

Baser, Derya., Kopcha, T. J., & Ozden, M. Y. (2015). Developing a technological pedagogical content knowledge (TPACK) assessment for preservice teachers learning to teach English as a foreign language. *Computer Assisted Language Learning*, *29(4),* 749–764.

Daryanto. (2005). *Evaluasi Pendidikan.* Jakarta: Rineka Cipta.

Desmita, Yulia., Suarman., & Gimin. (2021). Pedagogic and Professional Competencies of Social Science Subject Teachers in Relation to Motivation and Learning Achievement. *Journal of Educational Sciences*, *5(1),* 188-197.

Fisher Jr., W. P. (2007). Rating Scale Instrument Quality Criteria. Rasch Measurement Transaction, 21, p.1095. https://www.rasch.org/rmt/rmt211m.htm

Hayati, S., & Lailatussaadah. (2016). Validitas dan Reliabilitas Instrumen Pengetahuan Pembelajaran Aktif, Kreatif dan Menyenangkan (PAKEM) Menggunakan Model Rasch. *Jurnal Ilmiah DIDAKTIKA*, *16(2),* 169-179.

Lidinillah, D. A. M., Aprilia., Suryana, D., Ahmad., A. B. (2020). Development of Creativity Instrument through Rasch Model Analysis. *Universal Journal of Educational Research 8(4),* 1620-1627.

Majid, A. (2014). *Penilaian hasil belajar*. Jakarta: Bumi Aksara

Marina, Indrawati, Henny., & Suarman. (2019). Application of Moving Class Learning Models and Teacher Pedagogical Competence on Learning Motivation and Student Learning Discipline. *Journal of Educational Sciences*, *3(1),* 72-83

Muslich, M. (2011). *Penilaian berbasis Kelas dan Kompetensi*. Refika Aditama. Bandung

Nofrion, Wijayanto, B., Wilis, R., and Novio, Rery. (2018). Analisis *Technological Pedagogical and Content Knowledge* (TPACK) Guru Geografi di Kabupaten Solok, Sumatera Barat. *Jurnal Geografi. 10(2),* 105-116.

Plomp, T., & Nieveen, N. (2013). *Educational Design Research dalam Educational Design Research: An Introduction*. Enschede, University Of Twente.

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Yogyakarta: Trim Komunikata. ISBN 978-602-14371-1-7

Suryawati, E., & Roza, Yenita. (2017). Strengthening Prospective Teacher of Biology on Technological Pedagogical Content Knowledge (TPACK*). Applied Science and Technology*, *1(1)* 2017 http://www.estech.org

Sulistiawan, H. (2016). Kualitas Soal Ujian Sekolah Matematika Program IPA dan Kontribusinya Terhadap Hasil Ujian Nasional. *Jurnal Penelitian dan Evaluasi Pendidikan, 20(1),* 1-10

Yeh, Yi-Fen., Hsu, Ying-Shao., Wu,Hsin-Kai., Hwang, Fu-Kwun., Lin, Tzu-Chiang. (2014). Developing and validating technological pedagogical content knowledge- practical (TPACK-practical) through the Delphi survey technique. *British Journal of Educational Technology*, *45(4),* 707–722.

Yulvisriani, Suryawati, Evi., & Suzanti, Fitra. (2020). The Role of Empowerment of Dumai Science Teachers Association (MGMP IPA) in improving the Professional Competency of Junior High School Science Teacher. *Journal of Educational Sciences, 4(4),* 821-833.

How to cite this article:

Arnan, R. H., Suryawati, E., & Mahadi, I. (2021). Development of Knowledge Test Instruments of Prospective Bachelor of Biological Education Based on TPACK and KKNI Framework. *Journal of Educational Sciences*, *5(2)*, 224-234.