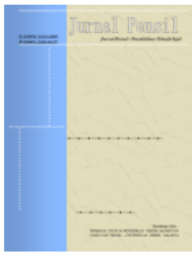
Available *online* at: http://journal.unj.ac.id

Jurnal Pensil | Pendidikan Teknik Sipil

Journal homepage: http://journal.unj.ac.id/unj/index.php/jpensil/index



THE DEVELOPMENT OF HOTS-BASED TECHNICAL MECHANICAL SUBJECT INSTRUMENTS USING A CLASSICAL TEST THEORY APPROACH

Salma Maharani[1], Zidni Husnurofik[2], Eva Maria Sianturi[3], Awaluddin Tjalla[4] Lussy Dwiutami Wahyuni[5], Ninda Ayu Narassati[6]*

[1,2,3,4,5] Program Studi Penelitian dan Evaluasi Pendidikan, Fakultas Pascasarjana, Universitas Negeri Jakarta
Jalan Rawamangun Muka, Rawamangun, Jakarta Timur, DKI Jakarta, 13220, Indonesia
[6] Program Studi Pendidikan Teknik Bangunan, Fakultas Teknik, Universitas Negeri Jakarta
Jalan Rawamangun Muka, Rawamangun, Jakarta Timur, DKI Jakarta, 13220, Indonesia
[*1]salma.maharani@mhs.unj.ac.id, [2]zidni.husnurofik@mhs.unj.ac.id,
[3]eva.maria.sianturi@mhs.unj.ac.id, [4]Awaluddin-Tjalla@unj.ac.id, [5]lussysf@unj.ac.id,
[6]nindaayunarassati@gmail.com

P-ISSN: 2301-8437
E-ISSN: 2623-1085

ARTICLE HISTORY
Accepted:
20 November 2023
Revision:
22 Januari 2024
Published:
30 Januari 2024

ARTICLE DOI:
10.21009/jpensil.v13i1.40225

Abstract

Along with the rapid development of the construction world, the competitiveness to work in this field has become strict so that graduates must have their own differentiators and attractions, one of which is the ability to think higher-level. In SMK in the field of building, one of the subjects that requires understanding concepts in solving problems is engineering mechanics. This subject is famous as a subject that applies derived concepts of mathematics, physics, and logic of critical thinking so the development of this instrument is intended to be able to help assess students' higher-order thinking skills. This study uses Research and Development Research Methods using the Classical Test Theory approach where after being tested on 574 respondents of vocational students in several areas in Jakarta, it was found that there were 19 questions with moderate reliability results and 1 question with low reliability, 19 questions with medium difficulty and 1 question with easy difficulty, and 19 questions with medium difference power and 1 question with low difference power.

Keywords: Classical Theory Test, Cognitive Instrument, Technical-Mechanical

Introduction

The development of learning in the 21st century requires teachers to be able to implement the Minimum Competency Assessment (AKM) program (Asmarni & Zakir, 2023) to support the development of 4C (Critical thinking, Communication, Collaboration, and Creativity) skills in students (Isrokatun et al., 2022; Ningsih et al., 2023; Widiyanto & Desstya, 2023) which is important for the learning process in this era. Teachers are expected to be able to develop their creativity and integrate the concept of Higher Order Thinking Skills (HOTS) into learning (Hanik et al., 2020) and evaluation oin the classroom (Riadi, 2017). Evaluation is a crucial process in a learning process, this process allows teachers to assess the effectiveness of the learning they have done and provide feedback to students (Hidma et al., 2023). According to existing research, teachers' ability to compile test instruments is still low (Hasanah et al., 2023; Sijabat, 2014). So far the test that made for students still tend to measure the ability to remember information, rather than measuring higher-order thinking skills (Saraswati & Agustika, 2020). This concept is no longer relate with what is needed in this era where it can hinder students' ability to develop critical, analytical, and creative thinking skills (Iskandar, 2013; Kamilati, 2018).

The vocational level is a level that is expected to be able to develop superior and competent human resource candidates and produce individuals who are ready to work after graduation in their vocational fields (Mardhiyah et al., 2021; Ratnata, 2012). SMK in the building sector is a secondary vocational education level that focuses on creating superior human resources to be ready to contribute to the civil sector (Blima et al., 2016; Ramadhan et al., 2013) Nowadays, SMK not only focuses on improving hard skills but also on improving soft skills, especially those related to higher-order thinking, but currently it is known that the ability of SMK students in this regard is still low. While, along with the rapid development of the construction world, the competitiveness to work in this field has become tight so that graduates must have a differentiator that can be a special attraction (Blima et al., 2016) the crucial thing is related to higher-order thinking skills and critical thinking power in solving problems (Arthur, 2019; Papantoniou, 2017).

In vocational education, one of the subjects that prioritizes understanding concepts and supports the improvement of higher-order thinking skills at the vocational level in the building sector is Engineering Mechanics (Maulana et al., 2022). This subject requires students to be able to understand the concepts and patterns of problem solving through logical calculations (Wardoyo & Ma'arif, 2015) where the concepts instilled in this subject are derivatives of the fields of mathematics and physics which are very suitable in terms of improving critical reasoning skills (Prajaka & Purwadi, 2016). However, teachers have not done much development of measuring instruments to measure the level of their students' ability to reason critically (Laksono, 2018) especially in this subject (Arthur et al., 2021) because the preparation of measuring instruments is only limited to measuring student memory. So through this research in order to support the improvement of student quality and help teachers to provide an overview of the formation of measuring instruments expected in 21st century learning in SMK in the building or related fields, a development of measuring instruments in the cognitive realm based on HOTS was made.

Research Methodology

This study uses RnD (Research and Development) research method with ADDIE development model. This development model consists of the stages of analysis, design and development, and implementation and evaluation (Sandra et al., 2022). The flow of instrument development is depicted in the flow chart in figure 1.

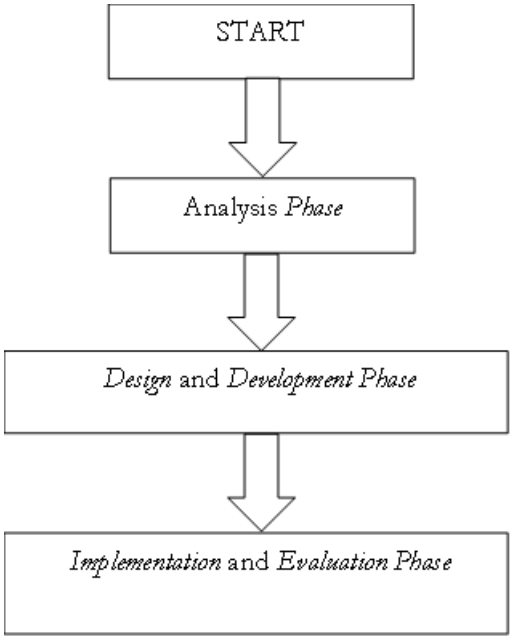


Figure 1. The Development Flow

At the analysis stage, an analysis of the Basic Competencies used when formulating Competency Achievement Indicators and Learning Objectives is carried out based on the syllabus of Engineering Mechanics Subjects. Then at the *design* and development *stage, the preparation of question* grids, selection of subject matter related to KD, formulation of question indicators, determination of cognitive levels of question items, determination of stimuli to be used in questions, writing multiple-choice question items of the HOTS type according to the grids, making scoring guidelines, and expert validation. Then at the implementation and *evaluation stage, the* instrument items were tested on students in several vocational schools in Jakarta, the results were processed using JMetrik software with a classical test theory approach to analyze each question item. The instruments developed in this study have been validated by 9 experts with details of 3 experts each for aspects of language, material, and instruments. In this study, the analysis that was emphasized was on the quality aspect of the instrument items distributed.

Research Results and Discussion

The table below containing the question instrument that has been tested for students of SMK in Building Sector.

Table 1. Instrument's Item

No.	Questions
1.	Look at the following picture! A ventilation canopy is given two styles P1 and P2. Where does the P1 force rotate to produce the moment?
2.	Look at the following picture! A potted plant is placed in the center of the concrete canopy as illustrated into a CD console beam. What is the condition of the beam after being loaded?
3.	Look at the following picture! The beam structure in the basic food building bears an even load q kN/m as the AB beam. What is the condition of the beam after being given a load?
4.	Look at the following picture! A hanging concrete canopy is given vertical loads of 100 kN and 60 kN down. What is the amount of vertical reaction on the canopy pedestal so that the structure is balanced?
5.	Look at the following picture! A balcony in a dwelling bears an even load q = 20 kN/m. The statement below that is exactly related to the description of the conversion formula Q is ...

No.	Questions
6.	Look at the following picture! The roof canopy of the house is given P1 and P2 styles. If Sin 37° = 0.6 and Cos 37° = 0.8. What is the normal force value of HF required for the structure to remain balanced?
7.	Look at the following picture! A pinch beam bears an even load and an oblique centered load. If Sin 53° = 0.8 and Cos 53° = 0.6, then the description of the equation to calculate the exact value of the vertical fulcrum reaction at point P is...
8.	Look at the following picture! One of the beams in the construction of the library building bears the load due to cabinets of q kN/m and chairs of P kN. The description of the moment equilibrium equation at the exact point K is...
9.	Look at the following picture! A canopy above the window is linked with a rope at the end illustrated as an HI console beam. The pulling force caused by the rope is 30 kN. If Sin 37° = 0.6 and Cos 37° = 0.8. What is the Normal Force (Nx) value generated by the HI console beam?
10.	Look at the following picture! The beam in a factory carries the load caused by a lifting machine of 10 kN/m and a production machine of 22 kN. If the value of the vertical force reaction VR = 10.9 kN and VS = -18.9 kN. What is the value of Shear Force before force P?
11.	Look at the following picture! The beams in the textile factory building bear two centralized loads in opposite directions as a result of the work of the machine. If the reaction of the vertical force fulcrum T = 7 kN and pedestal U = 5 kN. So what is the value of the Bending Moment that occurs at a distance of 6 m in terms of the pedestal T?
12.	Look at the following picture! Which is the correct statement below?
13.	Look at the following picture! The beam of a storage shed bears an oblique force P. If Sin 53° = 0.8 and Cos 53° = 0.6. What is the proper interpretation of the normal force occurring at a distance of 0 m to 1.4 m in view of the fulcrum G?
14.	Look at the following picture! A KL beam in the construction of the library building bears the load caused by bookshelves of q = 2 kN/m. If the reaction value of the fulcrum force VK = 7.5 kN and VL = 2.5 kN. What is the correct interpretation of the bending moment value at distance L/2?
15.	Look at the following picture! Structural beams in apartment buildings carry an even load of 30 kN/m. If VM = VN = 22.5 kN. What are the values of the shear force and bending moment at the distance L/2 respectively?
16.	Look at the following picture! A pinch beam accepts an inclined transverse force of 14 kN. If Sin 53° = 0.8 and Cos 53° = 0.6. Which is the right normal style diagram image?
17.	Look at the following picture! The above shear force diagram image is designed by a two-pedestal beam that has a load category...
18.	Look at the following picture! The concrete canopy is given a centralized load in the opposite direction. The design of the right shear force diagram drawing for the loading of the beam is...
19.	Look at the following picture! The beams on the building of the house bear two centralized loads in opposite directions. The design of the diagram drawing of the right moment for the loading of the beam is...
20.	Look at the following picture! In factory construction, the lifting machine load of 3 kN/m and the press of 4 kN/m produce VR force reaction = -4.25 kN and VS = 2.25 kN. The design of the right shear force diagram drawing for the loading of the beam is…

After analysis through JMetrik software, the results were obtained as follows. Based on the results of data processing that has been carried out on a total of 20 question items on engineering mechanics to 574 respondents, the analysis will be divided into analysis of item difficulty, differentiation, test level, and reliability.

The first analysis is related to the difficulty of the question items. In the classical test theory approach, the level of difficulty of question items is needed to provide categorization of the question items that have been made (Kholis, 2017; Prihatiningih et al., 2021) To determine the level of difficulty of the question items can be seen through the table below.

Table 2. Question Item Difficulty Category (Bagiyono, 2017)

P Value	Difficulty Category
0	Very difficult
$0 < P \leq 0.3$	Difficult
$0.3 < P \leq 0.7$	Moderate
$0.7 < P < 1$	Easy
1	Very Easy

In question number 1 obtained a value of 0.6742 with a medium difficulty category, question number 2 obtained a value of 0.6585 which received a medium difficulty category, then question number 3 with a difficulty value of 0.7143 which was included in the easy category, question number 4 with a score of 0.6655 was included in the medium category, question number 5 obtained a value of 0.4094 in the medium category, Question number 6 gets the medium category with a score of 0.6498, question number 7 gets the medium category of 0.5819, question number 8 gets the medium category with a gain of 0.6446, question number 9 gets the medium category with a gain of 0.3467, question number 10 gets the medium category with a gain of 0.6533, question number 11 gets the medium category with a gain of 0.6359, Question number 12 gets the medium category with a gain of 0.6481, question number 13 gets the medium category with a gain of 0.5784, question number 14 gets the medium category with a gain of 0.5941, question number 15 gets the medium category with a gain of 0.5732, question number 16 gets the medium category with a gain of 0.6167, question number 17 gets the medium category with a gain of 0.6202, Question number 18 gets the medium category with a gain of 0.6638, question number 19 gets the medium category with a gain of 0.5540, and question number 20 gets the medium category with a gain of 0.4791. Based on these results, 95% of all question items have a medium level of difficulty with 5% of all question items having easy categories. This indicates that out of 95% of all question items, the correct answer value given by students is in the middle, while the other 5% of question items students respond to answers correctly.

Based on the theory proposed by Anderson and Kratwohl (2001), a good question is a problem that is neither too easy nor too difficult. This is because questions that have a difficult level of difficulty will cause students to feel desperate to do it, while questions that have an easy level of difficulty cannot spur students to think deeper (Fatimah &; Alfath, 2019). So that the result of a moderate difficulty of 95% indicates that the quality of the questions is good in line with the theory proposed by Anderson and Kratwohl. The follow-up of the difficulty level of this question item can be used as a standard reference for making better HOTS questions, especially in the field of Engineering Mechanics, not only that in the preparation of difficulty level question items can also be used to measure the competence of teachers in making standardized questions to create more competent students (Fitrianawati, 2015).

Table 3. Discriminant Test Result

Item	Option (Score)	Difficulty	Std. Dev.	Discrimin.
1	Overall	0.6742	0.4691	0.6428
2	Overall	0.6585	0.4746	0.6795
3	Overall	0.7143	0.4521	0.6175
4	Overall	0.6655	0.4722	0.6710
5	Overall	0.4094	0.4922	0.1475
6	Overall	0.6498	0.4774	0.6650
7	Overall	0.5819	0.4937	0.7061
8	Overall	0.6446	0.4791	0.5870
9	Overall	0.3467	0.4763	0.2074
10	Overall	0.6533	0.4763	0.6462

Item	Option (Score)	Difficulty	Std. Dev.	Discrimin.
11	Overall	0.6359	0.4816	0.6802
12	Overall	0.6481	0.4780	0.7191
13	Overall	0.5784	0.4942	0.6445
14	Overall	0.5941	0.4915	0.6584
15	Overall	0.5732	0.4950	0.7055
16	Overall	0.6167	0.4866	0.5251
17	Overall	0.6202	0.4858	0.6726
18	Overall	0.6638	0.4728	0.4688
19	Overall	0.5540	0.4975	0.5303
20	Overall	0.4791	0.5000	0.4521

Next, based on the table above, different strengths or discriminants are also known. The discriminating power serves to provide an idea of how capable the problem is to be able to distinguish the abilities of students belonging to the upper and lower classes (Andriani & Siswanto, 2022; Ndiung & Jediut, 2020). The categorization of the values of each item can be seen based on the differentiating power category table below.

Table 4. Discriminant Category (Bagiyono, 2017)

D Value	Discriminant Category
$D \leq 0$	Very low
$0 < D \leq 0.2$	Low
$0.2 < D \leq 0.4$	Modearte
$0.4 < D \leq 0.7$	High
$0.7 < D \leq 1$	Very High

For the difference between question number 1 obtained a high category of 0.6428, question number 2 obtained a high category of 0.6795, question number 3 obtained a high category of 0.6175, question number 4 obtained a value of 0.6710 with a high category, question number 5 obtained a value of 0.1475 with a low category, question number 6 obtained a value of 0.6650 with a high category, Question number 7 obtained a score of 0.7061 with a very high category, question number 8 obtained a score of 0.5870 with a very high category, question number 9 obtained a score of 0.2074 with a medium category, question number 10 obtained a score of 0.6462 with a high category, question number 11 obtained a score of 0.6802 with a high category, question number 12 obtained a score of 0.7191 with a very high category, Question number 13 obtained a score of 0.6445 with a high category, question number 14 obtained a score of 0.6584 with a high category, question number 15 obtained a score of 0.7055 with a very high category, question number 16 obtained a score of 0.5251 with a high category, question number 17 obtained a score of 0.6726 with a high category, question number 18 obtained a score of 0.4688 with a high category, Question number 19 gets a score of 0.5303 in the high category, and question number 20 gets a score of 0.4521 in the very high category. That is, of all 20 items in question item number 5 get low discriminating power so that it can be a special consideration for instrument makers whether to be corrected or deleted because this item cannot distinguish the abilities of students while the other 20 have received very high value categories, meaning that the question already has good quality because the value of the distinguishing power of the question is directly proportional to quality (Fajar & Junaidi, 2022). Point number 9 is also the only item that has distinguishing power from the medium category, meaning that this item does not need to be deleted but makers can consider revising or changing the question item (Nurmaliza, 2022). For other items, it is included in the high category so that the items are considered to be able to distinguish students' abilities (Solichin, 2017).

Table 5. Level Statistics Test Result

Test Level Statistics	
Number of Items	20
Number of Examinees	574
Min	0.0000
Max	20.0000
Mean	11.9617
Median	14.0000
Standard Deviation	6.0763
Interquartile Range	11.0000
Skewness	-0.3544
Kurtosis	-1.4044
KR21	0.9156

Next is about the results of statistical level tests. Of the 20 question items with 574 respondents, they get an average of 11.9617 with a middle value of 14.0000, a standard deviation of 6.0763, an interquartile value of 11.0000, a skewness value of -0.3544 which means that if made into a graph, the graph will tilt to the left of 0.3544, the kurtosis value of -1.4044 which means that the data distribution is relatively flat, and for the KR21 result of 0.9156 which means it has a high reliability value especially For exact based subjects that have higher standards of reliability values compared to social-based subjects.

The next analysis is related to reliability. Reliability assessments are divided into test scales and item scales. Below is a reliability analysis according to a test scale consisting of the Guttman scale, Cronbach Alpha Coefficient, Feldt-Gilmer, Feldt-Brennan, and Raju's Beta.

Reliability Analysis

Reliability per test scale is the result of calculations from several commonly applied methods related to reliability calculations. The results can be seen through the table 6.

Table 6. Reliability Analysis per Test Scale

Method	Estimate	95% Conf. Int.	SEM
Guttman's L2	0.9239	(0.9145, 0.9326)	1.6782
Coefficient Alpha	0.9201	(0.9104, 0.9293)	1.7185
Feldt-Gilmer	0.9236	(0.9142, 0.9324)	1.6813
Feldt-Brennan	0.9226	(0.9131, 0.9315)	1.6916
Raju's Beta	0.9201	(0.9104, 0.9293)	1.7185

Based on the table above, all the results of the reliability analysis above get a value above 0.9 which can be categorized as high reliability. The methods in the reliability analysis table can later be further specified to be used as a reference basis for future research or similar research. In addition to analysis by test scale, there are also reliability analysis results per item scale which can be seen in the table below if every one of the items is deleted/removed.

Reliability If Item Deleted

The table below displays the results of the reliability analysis per item if one of the questions is deleted. This is to show the level of influence of one question on the overall value of reliability. The goal, if there is a problem deleted and then produces a much higher reliability value than before, it can later boil down to making a decision whether the question will be replaced or deleted / not used.

Table 7. Reliability Analysis per Item

Item	L2	Alpha	F-G	F-B	Raju
1	0.9189	0.9149	0.9187	0.9176	0.9149
2	0.9182	0.9141	0.9178	0.9168	0.9141
3	0.9196	0.9155	0.9193	0.9182	0.9155
4	0.9184	0.9143	0.9181	0.9170	0.9143
5	0.9281	0.9257	0.9280	0.9274	0.9257
6	0.9184	0.9144	0.9181	0.9170	0.9144
7	0.9175	0.9134	0.9171	0.9160	0.9134
8	0.9202	0.9161	0.9199	0.9189	0.9161
9	0.9269	0.9242	0.9269	0.9261	0.9242
10	0.9189	0.9148	0.9186	0.9175	0.9148
11	0.9181	0.9140	0.9178	0.9167	0.9140
12	0.9172	0.9132	0.9169	0.9158	0.9132
13	0.9189	0.9148	0.9186	0.9175	0.9148
14	0.9184	0.9145	0.9182	0.9171	0.9145
15	0.9174	0.9134	0.9171	0.9160	0.9134
16	0.9214	0.9175	0.9211	0.9201	0.9175
17	0.9183	0.9142	0.9180	0.9168	0.9142
18	0.9225	0.9187	0.9222	0.9212	0.9187
19	0.9215	0.9174	0.9213	0.9202	0.9174
20	0.9231	0.9192	0.9229	0.9218	0.9192

L2 : Guttman's lamba-2
Alpha: Coefficient alpha
F-G: Fedlt-Gilmer coefficient
F-B: Fedlt-Brennan coefficient
Raju: Raju's beta coefficient

In questions number 1, 2, 3, and 4 if the item is deleted / omitted, it is found that on the L2, Alpha, F-G, F-B, and Raju test scales have decreased, this gives a slight change in the average reliability value of the whole meaning that the difference in the question can be interpreted moderately because the effect of the decrease given is not too significant. Question number 5, if the item is deleted/omitted, it is found that on the L2, Alpha, F-G, F-B, and Raju test scales have increased, this gives a slight change in the average reliability value of the whole meaning that the difference in the question can be interpreted low. For questions number 6 – 20 have the same interpretation results as questions number 1 – 4 where if the item is deleted / omitted, the results are obtained that on the L2, Alpha, F-G, F-B, and Raju test scales have decreased which gives a slight change in the average reliability value of the whole meaning that the difference in the question can be interpreted moderately.

Conclusion

This research uses development and research methods where the instrument validation involves 3 experts including linguists, materials, and instruments as well as trials to as many as 574 respondents. The trial data that has been analyzed using the Jmetric application obtained the results of the analysis for the difficulty level of medium question items for question items number 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20 while there is one question that is relatively easy, namely question number 5. For the difference in points, questions number 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20 are included in the medium category and for question number 5 are included in the easy category. For overall reliability results, the question items are already of high value, if items 1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, and 20 are deleted causing the reliability results to decrease slightly which means confirming that the items

have moderate difference power while if item number 5 is deleted it increases the reliability value which confirms that the difference power of the question is low.

References

Anderson, L. W., & Krathwohl, D. R. (2001). A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Education Objectives. Addison-Wesley Lonman Inc.

Andriani, D. I. A., & Siswanto. (2022). Analisis Butir Soal Mata Pelajaran Akuntansi Dasar Kelas X Akuntansi. Kajian Pendidikan Akuntansi Indonesia, 11(5), 37–52.

Arthur, R. (2019). Keakurasian Penyetaraan Vertikal dengan Metode Ekuipresentil pada Kemampuan Berpikir Tingkat Tinggi (HOTS). Jurnal Pendidikan Edutama, 6(2), 1. https://doi.org/10.30734/jpe.v6i2.156

Arthur, R., Daryati, Maulana, A., Febiansyah, R., & Kidung, K. (2021). Scientific Literacy of Vocational School Students in Building Construction. Journal of Physics: Conference Series, 1833(1). https://doi.org/10.1088/1742-6596/1833/1/012036

Asmarni, T., & Zakir, S. (2023). Persepsi Guru MTsN 1 Lima Puluh Kota tentang Asesmen Kompetensi Minimum. Arus Jurnal Pendidikan (AJUP), 3(1), 6–14.

Bagiyono. (2017). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1. Widyanuklida, 16(1), 1–12. https://doi.org/10.2307/40202478

Blima, O., Ahmad, D., & Nindyawati. (2016). Meningkatkan Technical Skill Siswa SMK Teknik Bangunan Melalui Pelaksanaan Praktik Kerja Industri. Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan, 1(4), 681–685.

Fajar, R., & Junaidi, J. (2022). Analisis Butir Soal Tipe Multiple Choices Questions (MCQ) pada Ujian Akhir Semester Ganjil Mata Pelajaran Sosiologi Kelas XII SMA. Jurnal Sikola: Jurnal Kajian Pendidikan Dan Pembelajaran, 4(1), 23–34. https://doi.org/10.24036/sikola.v4i1.194

Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda, dan Fungsi Distraktor. AL-MANAR: Jurnal Komunikasi Dan Pendidikan Islam, 8(2), 37–64.

Fitrianawati, M. (2015). Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru dan Hasil Belajar Peserta Didik. Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa, 5(3), 282–295.

Hanik, U., Mardiyana, I. I., & Rosid, A. (2020). Peningkatan Kompetensi Guru dalam Mengintegrasikan HOTS (Higher Order Thinking Skills) Pada Pembelajaran. Jurnal Ilmiah Pangabdhi, 6(1), 44–48. https://doi.org/10.21107/pangabdhi.v6i1.7101

Hasanah, M. N., Nur'aini, H. D., Aliyah, H., Aji, J. L. F. P., Azuma, M. A., Khoiriyah, M., & Salsabila, R. (2023). Identifikasi Kesulitan Guru dalam Menyusun Instrumen Penilaian Anak Berkebutuhan Khusus. Jurnal Penelitian Pendidikan Kebutuhan Khusus, 11(2), 62–70.

Hidma, C. A., Livinti, L., Afany, S., & Syafiq, Z. Z. (2023). Peran Guru dalam Evaluasi Pembelajaran di Sekolah. Jurnal Motivasi Pendidikan Dan Bahasa, 2(1), 123–132.

Iskandar, A. (2013). Pengembangan Perangkat Penilaian Psikomotor di Sekolah Menengah Kejuruan (SMK). Inspiration: Jurnal Teknologi Informasi Dan Komunikasi, 3(1), 37–46.

Isrokatun, I., Yulianti, U., & Nurfitriyana, Y. (2022). Analisis Profesionalisme Guru dalam Pelaksanaan Pembelajaran Daring di Masa Pandemi Covid-19. Jurnal Basicedu, 6(1), 454–462. https://doi.org/10.31004/basicedu.v6i1.1961

Kamilati, N. (2018). Analisis Komponen Penilaian Pada Rencana Pelaksanaan Pembelajaran Sebagai Acuan Pengembangan Kurikulum Diklat Teknis Substantif Guru. EDUKASI: Jurnal Penelitian Pendidikan Agama Dan Keagamaan, 16(1), 1–17. https://doi.org/10.32729/edukasi.v16i1.440

Kholis, R. A. N. (2017). Analisis Tingkat Kesulitan (Difficulty Level) Soal Pada Buku Sejarah Kebudayaan Islam Kelas 8 Kurikulum 2013. INTAJ : Jurnal Penelitian Ilmiah, 01(02), 93–113. https://doi.org/10.35897/intaj.v1i2.111

Laksono, P. J. (2018). Pengembangan dan Penggunaan Instrumen Two-Tier Multiple Choice pada Materi Termokimia Untuk Mengukur Kemampuan Berpikir Kritis. Orbital: Jurnal Pendidikan Kimia, 2(2), 80–92. https://doi.org/10.19109/ojpk.v2i2.2646

Mardhiyah, R. H., Aldriani, S. N. F., Chitta, F., & Zulfikar, M. R. (2021). Pentingnya Keterampilan Belajar di Abad 21 Sebagai Tuntutan dalam Pengembangan Sumber Daya Manusia. Lectura: Jurnal Pendidikan, 12(1), 29–40.

Maulana, A., Puspita, A. J., Pangastuti, K. K., Daryati, & Arthur, R. (2022). The Concept of Literacy Vocational-Based E-Module of Technical Mechanical Subject. Journal of Physics: Conference Series, 2377, 1–5. https://doi.org/10.1088/1742-6596/2377/1/012068

Ndiung, S., & Jediut, M. (2020). Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar BBerorientasi pada Berpikir Tingkat Tinggi. Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran, 10(1), 94. https://doi.org/10.25273/pe.v10i1.6274

Ningsih, P. R., Dellia, P., Risnasari, M., Cahyani, L., Rikanawati, R., & Albian, R. (2023). Pengembangan Soal Asesmen Kompetensi Minimum (AKM) Pada Materi Sistem Persamaan Linear Dua Variabel (SPLDV) dengan Media Live Worksheet. Jurnal Ilmiah Edutic : Pendidikan Dan Informatika, 9(2), 178–187. https://doi.org/10.21107/edutic.v9i2.17885

Nurmaliza. (2022). Analisis Kemampuan Berpikir Kritis Matematis Siswa Ditinjau Dari Habits Of Mind Siswa SMAN 1 Tembilahan Hulu. Universitas Islam Riau.

Papantoniou, E. (2017). WEB based technical problem solving for enhancing writing skills of secondary vocational students. Education and Information Technologies, 22(4), 1825–1852. https://doi.org/10.1007/s10639-016-9520-y

Prajaka, H., & Purwadi, D. (2016). Hubungan Penguasaan Matematika dan Fisika Terhadap Penguasaan Mekanika Teknik Pada Siswa SMK Negeri di Surabaya. Jurnal Kajian Pendidikan Teknik Bangunan, 2(2), 234–240.

Prihatiningih, N., Awalia, P. J., Oktavia, I., Walid, A., Kusuma, R. G. T., & Hadiwinarto. (2021). Analisis Perbandingan Tingkat Kesulitan Butir Soal Ujicoba UN-IPA dari SMPN 05 Kota Bengkulu dan Dinas Pendidikan Kota Bengkulu Tahun 2019/2020. Eduproxima: Jurnal Ilmiah Pendidikan IPA, 3(2), 72–76.

Ramadhan, M. A., Iriani, T., & Handoyo, S. S. (2013). Relevansi Kompetensi Lulusan SMK Khususnya Kompetensi Keahlian Teknik Gambar Bangunan dengan Kompetensi yang Dibutuhkan di Dunia Kerja. Jurnal PenSil, 2(1), 1–10. https://doi.org/10.21009/jpensil.v2i1.7282

Ratnata, I. W. (2012). Konsep Pemikiran Dalam Pengembangan Pendidikan Vokasi Untuk Menghadapi Tuntutan Dunia Kerja. Seminar Internasional Aptekindo, 41–46.

Riadi, A. (2017). Kompetensi Guru dalam Pelaksanaan Evaluasi Pembelajaran. Ittihad Jurnal Kopertais Wilayah XI Kalimantan, 15(28), 52–67. https://doi.org/10.30998/rdje.v6i1.4371

Sandra, R. O., Maison, & Kurniawan, D. A. (2022). Pengembangan Instrument Miskonsepsi Menggunakan Dreamweaver Berbasis Web pada Materi Tekanan Berformat Five-Tier. Jurnal Fisika : Fisika Sains Dan Aplikasinya, 7(1), 22–28. https://doi.org/10.35508/fisa.v7i1.6575

Saraswati, P. M. S., & Agustika, G. N. S. (2020). Kemampuan Berpikir Tingkat Tinggi Dalam Menyelesaikan Soal HOTS Mata Pelajaran Matematika. Jurnal Ilmiah Sekolah Dasar, 4(2), 257–269. https://doi.org/10.23887/jisd.v4i2.25336

Sijabat, L. (2014). Meningkatkan Kompetensi Guru Menyusun Instrumen Tes Melalui Model Pelatihan SSOTT di SMA dan SMK Kabupaten Dairi. Jurnal Pendidikan Dan Kepengawasan, 1(1), 107–118.

Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan Dalam Evaluasi Pendidikan. Dirasat: Jurnal Manajemen Dan Pendidikan Islam, 2(2), 192–213. https://doi.org/https://doi.org/10.26594/dirasat.v2i2.879

Wardoyo, T. C. T., & Ma'arif, F. (2015). Pengembangan Media Pembelajaran Berbasis Video Animasi Pada Mata Pelajaran Mekanika Teknik di SMK Negeri 1 Purworejo. Jurnal Elektronik Mahasiswa Pendidikan Teknik Sipil (JEPTS), 3(3), 1–7.

Widiyanto, R. R., & Desstya, A. (2023). Evaluasi Pelaksanaan Asesmen Kompetensi Minimum (AKM) pada Kompetensi Dasar Literasi Membaca Peserta Didik Sekolah Dasar. Jurnal Inovasi Pendidikan Dan Pembelajaran Sekolah Dasar, 7(2), 296–309. https://doi.org/10.31004/basicedu.v6i3.2907