

Content Validity by Experts Judgment Thermal and Transport Concept Inventory (TTCI) Assessment Integrated STEM to Measure Student's Problem-Solving Skills

Erlina Fatkur Rohmah^{1*}, Sukarmin², Daru Wahyuningsih³
^{1,2,3}Sebelas Maret University Indonesia

ABSTRACT

The study aimed to analyze the content validation of the STEM-integrated on thermal and transport concept inventory instrument used to measure the problem-solving abilities of high school students. The instrument questions developed amounted to nine description questions. This type of study is development research. The steps in this research are define, design, development, and dissemination. However, in this study, it was limited to the design stage (expert validation). Furthermore, the instrument was assessed qualitatively and quantitatively by a validator consisting of three experts and five users. Assessment aspects include material, language, and construction. The qualitative data collected was analyzed qualitatively, and the quantitative data were analyzed using the Content Validity Ratio (CVR). The results showed that nine items had a CVR value exceeding 0.75, which meant that all items had valid criteria for content validity. The CVR value obtained is between 0.88 and 0.96, which means that the instrument has high criterion validity.

Keywords: Content Validation, Assessment, Concept Inventory, STEM, Problem Solving.

INTRODUCTION

Assessment is an essential thing that happens in the world of education. Because assessment is the key to a teachers success in helping students succeed in learning (Sukenti dkk, 2020). Adeleke & Joshua 2015 stated that test results serve as guidelines in analyzing failures, evaluating learning outcomes, evaluating curriculum, and providing motivation to students to learn. The existence of an assessment carried out by the teacher will motivate students to study more diligently and achieve good learning outcomes (Mahirah, 2017).

Leaned assessment instruments play an essential role in improving teaching and curriculum (Laverty & Caballero, 2018). Bashoori & Supahar 2018 also said that the quality of learning can be increased by improving the assessment in measuring student learning outcomes. Ideally, assessments are carried out using principles, procedures, and instruments and paying attention to certain steps by considering time, place, and other factors (Herni et al, 2018). There-fore the teacher must consider the right test to measure the student's ability. Various types of tests in the assessment include multiple-choice, short-answer, and essay tests (Subali, 2016)). The description text uses to develop reasoning in solving student problems (Munadi, 2018).

A concept inventory is an assessment-based survey designed to test students understanding of specific physics concepts (Madsen, Mckagan, & Sayre, 2017). Conceptual understanding is the main content that students need to learn physics (Basin, Mađrala, Wařsik, & Małoszewski, 2017; Burkholder, Walsh, & Holmes, 2020; Kurniawati, Hartanto, & Zamzaili, 2017). Widodo (2006) stated that the right initial step in learning physics is to understand the concept first (Zhang & Zhang, 2018; Zheng et al., 2021). Understanding concepts is a

prerequisite for success in the study of physics as physics is not a subject to be memorized but requires reasoning and conceptual understanding (Linawati, 2018). Possessed students can achieve student success in learning physics (Puspitasari et al., 2021). However, student's conceptual understanding is still low and they are used to memorizing and memorizing formulas rather than understanding materials (Capriconia, 2022).

STEM education is an interdisciplinary approach that combines independent knowledge to enable students to solve real problems (Le et al, 2021). STEM education is believed to equip students with transdisciplinary knowledge and skills to deal with problems that occur in everyday life, including complex societies in the future (Le et al, 2021). However, there are still teachers who face many problems in implementing STEM education (Lee et al, 2019; Margot & Kettler, 2019; Ryu et al., 2019).

Corresponding Author e-mail: erlinafatkurrohmah@gmail.com

<https://orcid.org/0000-0003-3422-1912>

How to cite this article: Rohmah F E, Sukarmin, Wahyuningsih D (2023), Content Validity by Experts Judgment Thermal and Transport Concept Inventory (TTCI) Assessment Integrated STEM to Measure Student's Problem-Solving Skills, Vol. 14, No. 2, 2024, 316-323

Source of support: Nil

Conflict of interest: None.

DOI: 10.47750/pegegog.14.02.36

Received: 11.01.2023

Accepted: 30.07.2023

Publication: 01.04.2024

Problem-solving skills are a problem that is highlighted in student learning as a fundamental part of learning science in schools (Tong, Loc, Uyen, & Truc, 2020). Student's ability to understand can be seen from the test results in solving problems (Meiliani et al., 2021; Sulman et al., 2020; Zb et al., 2020), one of them is learning physics (Rozal et al., 2021; Sulman Sutopo, et al., 2021; Zb, Novalian, Ananda, et al., 2021). As it is known that physics is an empirical science, so when solving problems in physics, students must first understand the concept of the material under study. Students need to develop their thinking skills, not just memorize the lessons (Zb et al., 2020), but students must understand the concepts. They can solve problems and find solutions to problems. Jua, Sarwanto, & Sukarmin (2018) The percentage of students problem-solving skills based on the test shows that 52.93% indicators are different. These results show that students problem-solving abilities are in the lower category. It is in line with Azizah et al (2016) that the problem-solving abilities of students in Indonesia in learning physics are categorized as low. When students are low in solving problems, they may not understand the particular material being discussed (Batlolona, Baskar, Kurnaz, & Leasa, 2018).

The use of a good assessment instrument will produce good data. Based on the needs analysis that has been done, assessing students' problem abilities is very difficult and must be done objectively. Therefore, the instrument used must be able to measure accurately (Kurniawan et al, 2018). The right instrument will minimize measurement errors (Sinaga, 2016). Whether the test instrument is good or not is determined by its validity and reliability (Dewi & Sudaryanto, 2020). The validity of a good instrument by the Ministry of National Education (2017) according to the four aspects, namely; 1) material aspect, 2) construction aspect, 3) language aspect, and 4) higher order thinking ability aspect.

Jati & Aprilia (2019) shows that all teachers participated in socialization and training to develop test tools. However, in reality, there are teachers for various reasons in a school who have not been able to prepare assessment instruments properly and correctly (Husin, 2020). The survey results conducted with high school physics teachers showed that teachers rarely reviewed the final exam questions (Kaleka, 2016). Assessment of concept inventory ratings have not been tested for validity and reliability in most physical studies so the use of teacher-created instruments is unknown (Sukarmin, 2021). Some teachers only take the questions in the textbook and never do a qualitative review (Hamid, 2017). Even from the results of questionnaires and interviews, the assessment instruments made by the teacher did not assess the understanding of a material as a whole. The most common reason is that the teacher has difficulty sharing time in delivering material and conducting assessments (Sutama, S., Sandy, G. A., & Fuad, 2017).

Based on the information presented above, the Thermal and Transport Concept Inventory assessment instrument integrated with STEM to test problem-solving abilities needs to be analyzed for validity. The purpose of this study was to analyze the content validity of the scale. This study is to produce results that can be used as a basis for content validity in developing a STEM-integrated concept inventory

assessment instrument to measure problem-solving abilities.

Previous research on the content validity of the inventory concept only focused on quantitative analysis (Fatih, 2018; Mchrnaz, 2019; Zohre, 2017) while research on Rukman (2018) only focused on qualitative analysis. So, the novelty of this research is to analyze the validity of the contents of the inventory concept instrument using quantitative and qualitative analysis accompanied by pictures and data tables.

METHOD

Research Design

The type of research used is the development of a thermal and transport concept inventory (TTCI) assessment integrated stem to measure student's problem-solving skills. This research used a type of research and development (R&D) based on Thiagarajan & Semmel (Eny et al, 2021).

The development flow is shown in Figure 1

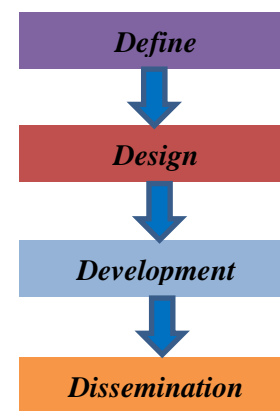


Figure 1: Development Flow

The research steps are described in the following steps :

Define

The study steps are described as follows. The definition stage begins with conducting a needs analysis by distributing questionnaires to 40 physics teachers in Indonesia and conducting interviews with ten physics teachers. Besides that, they also look for study material from instrument development study references.

Design

The needs analysis and study materials are developed up to the design stage. Data explained the development of an integrated STEM concept inventory assessment instrument on Thermal and Transport material to measure the problem-solving abilities of high school students. In designing a product, the steps taken are to determine core competencies, basic competencies, indicators of problem-solving abilities, indicators of concept inventory, indicators of questions, and assessment guidelines. The initial product design was evaluated by three experts and five users. After evaluation, the revisions were made based on the suggestions given.

Development

At the product assessment stage, trials were carried out in schools. The trials carried out included small-scale trials, large-scale trials, and operational trials.

Dissemination

The final stage is dissemination. the dissemination stage was carried out to obtain the teacher's response to the developed instrument. This response aims to measure the practicality of the instrument.

However, it was limited only to the design stage in the validation section.

Population and Sample/ Study Group/Participants

The needs analysis was carried out by 40 physics teachers in Indonesia, while the validation was by three experts and five users, namely high school physics teachers. It is recommended to have a minimum of six experts to get a loose CVR value (Hendryadi, 2017), but it is not recommended that experts rate more than ten (Puspitasari, 2021).

Data Collection Tools

Data validation results are qualitative data and quantitative data. Qualitative data were obtained from the validator's comments and suggestions for improving the products. Qualitative assessment of the validator makes improvements to produce a good quality instrument. Quantitative data is in the form of each validator's assessment score on product quality in terms of material, construction, and language.

Data Collection

Techniques for collecting data about subjects and measurements are questionnaires and suggestion questionnaires. the questionnaire is an instrument research that contains questions and will be answered by respondent (Sugiyono, 2014). The questionnaire used in this study has closed. The reason for using a closed questionnaire is that the researcher determines the statement answers, and the respondents choose their answers according to their actual opinions and circumstances. The dissemination technique in this study was carried out in two ways, namely offline and online through the google form platform.

Aspects of assessing the quality of the instrument can be seen in the following table:

Table 1: Aspect and Items

Aspect	Items
Material	Questions according to indicators of problem-solving ability
	The questions are following the concept of inventory indicators
	The questions contain the concepts of physics material that represent the ability to solve problems
	Contains clear instructions about the procedure for working on the questions
	The questions presented are in the field of high school physics
Language	The sentences used in the questions are Indonesian according to the EYD
	The sentences used are easy to understand
	The sentence used does not have a double meaning
Construction	The language used is communicative
	Statements encourage students to problem-solving skills
	Keep the topic statement concise
	The main questions are free from statements that have multiple meanings

The scoring criteria for each question are shown in the table below.

Table 2: Rating Scale

Symbol	Scale	Descriptor
TR	1	The items are completely irrelevant to the material, construction, and language aspects
KR	2	The items are less relevant to the material, construction, and language aspects
R	3	The items are relevant to the material, construction, and language aspects and are very relevant
SR	4	The items are very relevant to the material, construction, and language aspects

Data Analysis

The values obtained from each validator were analyzed using Aiken's formula (Aiken, 1980) calculations to determine the validity value of each item.

$$V = \frac{S}{(n(c-1))}$$

With, $S = \sum ni (r - l_0)$

The minimum CVR value according to Lawshe (1975) is presented in table 3 below

Table 3: minimum value CVR (Lawshe, 1975)

Number of rater	Minimum value
5	0.99
6	0.99
7	0.99

The validity of an assessment is determined by accepted standards for the interpretation of the assessment. The interpretation of the Aikens V formula is between 0 and 1 based on expert judgment (Sihole 2015). Standard verification see Table 4 below

Table 4: Validation criteria

Validation results	Validation Criteria
$0,8 < V \leq 1$	High
$0,4 < V \leq 0,8$	Medium
$V \leq 0,4$	Low

(Susangko, 2010)

FINDINGS

The content validity of the Thermal and Transport Concept Inventory assessment instrument was carried out by eight validators, namely three expert validators, and five user validators for 1) relevance of the item items to Thermal and Transport material, 2) evaluated for classification including relevance, 3) the relevance of the items to the concept inventory, 4) standardization of the language used, 5) functional descriptions and pictures.

Goodwin and Leech (2003) The meaning of the content test is based on the logical analysis and expert evaluation of the

measurement content such as items question types and sentences. Lynn (1986) states that the minimum number of experts used for expert verification is 3 and the maximum is 10.

The assessment process is conducted qualitatively and quantitatively using questionnaires that include core competencies indicators of conceptual items indicators of problem-solving level of awareness form of the item discussion questions evaluation and suggestions for improvement. The above sections are summarized in the questionnaire shown in Figure 2 below.

KARTU SOAL URAIAN








Identitas	BUTIR SOAL	PEMBAHASAN														
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td>Jenjang</td><td>: SMA</td></tr> <tr><td>Kelas</td><td>: XI</td></tr> <tr><td>Berusutan</td><td>: Erlina F</td></tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr><td>MATERI</td><td>NO SOAL</td></tr> <tr><td>Pemuaian</td><td>5</td></tr> <tr><td>INDKATOR CONCEPT INVENTORY</td><td>Level Kognitif</td></tr> <tr><td>Crosscutting conceent</td><td>C6</td></tr> </table> <p style="text-align: center;">INDIKATOR SOAL</p> <p>Menjelaskan penyebab rel kereta api bengkok.</p>	Jenjang	: SMA	Kelas	: XI	Berusutan	: Erlina F	MATERI	NO SOAL	Pemuaian	5	INDKATOR CONCEPT INVENTORY	Level Kognitif	Crosscutting conceent	C6	<p>Kondisi cuaca yang panas dapat berpengaruh pada rel kereta. Kondisi rel kereta api yang memuai atau melengkung terjadi di Porong, Sidoarjo senerti gambar berikut menjadi salah satu hal yang harus diperhatikan oleh operator kereta api.</p>  <p style="text-align: center;">Rel Kereta Api Benzok</p> <p>Pertanyaan :</p> <ol style="list-style-type: none"> Berdasarkan informasi di atas, permasalahan apa yang terjadi? Mengapa permasalahan tersebut dapat terjadi? Bagaimana upaya yang dilakukan untuk menghindari permasalahan tersebut? Kesimpulan apa yang dapat diambil berdasarkan permasalahan tersebut 	<ol style="list-style-type: none"> Permasalahan yang muncul adalah rel kereta api yang membengkok akibat cuaca yang panas. Hal tersebut diakibatkan oleh rel yang melengkung akibat pemuaian besi saat cuaca panas. Cara mengatasinya adalah dengan pemberian celah pada saat pemasangan rel kereta api. Kesimpulan dari permasalahan tersebut adalah besi memuai salah satu benda padat yang akan mengalami pemuaian ketika terkena suhu panas. Saat memuai besi pada rel kereta api akan menjadi melengkung. Untuk mengatasi permasalahan tersebut maka pemasangan rel kereta api harus diberi celah.
Jenjang	: SMA															
Kelas	: XI															
Berusutan	: Erlina F															
MATERI	NO SOAL															
Pemuaian	5															
INDKATOR CONCEPT INVENTORY	Level Kognitif															
Crosscutting conceent	C6															
<p style="text-align: center;">SARAN PERBAIKAN</p> <div style="border: 1px solid black; height: 50px; width: 100%;"></div>																
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th colspan="4" style="text-align: center;">VALIDASI</th> </tr> <tr> <th style="width: 25%;">TR</th> <th style="width: 25%;">KR</th> <th style="width: 25%;">R</th> <th style="width: 25%;">SR</th> </tr> </thead> <tbody> <tr> <td style="height: 20px;"></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>			VALIDASI				TR	KR	R	SR						
VALIDASI																
TR	KR	R	SR													

Figure 2: Sample Question Card Format

Qualitative assessment is carried out by providing a note of advice on the item items available on the question card. Notes and suggestions can be seen in the following table:

Table 5: Expert Advice and Comments

No.	Suggestion	First Picture	Image after repair
1.	Correct the questions according to the question indicators	<p>1. Ramadhan adalah bulan penuh berkah. Keseruan Ramadhan sangat terasa ketika menyiapkan menu untuk berbuka puasa. Menu berbuka puasa disajikan sayur asem, tempe, dan ayam. Adapun minumannya terdiri dari teh dan Teh panas sebagaimana diperlihatkan pada gambar di bawah. Teh untuk Dita dan adiknya sedangkan teh panas untuk ibu dan ayah. D bertugas untuk menyiapkan minumannya.</p>  <p>Es Teh Teh Panas</p> <p>Dita merasa kepanasan saat memegang teh panas dan saat memegang Teh, ia merasa dingin.</p> <p>Pertanyaan:</p> <ol style="list-style-type: none"> Tuliskan permasalahan yang dialami Dita! Apa penyebab Dita mengalami permasalahan tersebut? (Jelaskan secara konsep fisika)! 	<p>1. Lina berada di sebuah ruangan yang di dalamnya terdapat kursi. Saat bagian aluminium di pegang, Lina merasa dingin, namun saat bagian kayunya dipegang, Lina tidak merasakan dingin seperti pada gambar dibawah ini.</p>  <p>Memegang bagian aluminium Memegang bagian kayu</p> <p>Lina merasa heran, karena kursi tersebut berada dalam ruangan dengan suhu yang sama. Lina kemudian mengambil termogun dan mengukur suhu kedua bahan tersebut. Ternyata, suhu yang ditunjukkan oleh termogun tersebut sama.</p> <p>Pertanyaan:</p> <ol style="list-style-type: none"> Permasalahan apa yang sedang dihadapi Lina? Mengapa hal itu bisa terjadi?
2.	Fix pictures and stories on the problem	<p>1. Unit Layanan Pelanggan (ULP) PLN Sragen merencanakan tim untuk memperbaiki kabel yang terputus di simpang empat Bangsi, Kelurahan Sine, Sragen pada Rabu, 6 Oktober 2021 sebagaimana diperlihatkan pada gambar di bawah.</p>  <p>Gambar 3. Menepertaili Kabel Putus</p> <p>Kabel jaringan bertegangan 220 Volt terputus karena aksi tukang menyebarkan awan petir yang sedang berada di bawah memegangi stamper akibat serangan listrik. Manajer ULP PLN Sragen menjelaskan bahwa peristiwa tersebut diakibatkan oleh pemasangan kabel yang terlalu kencang apabila suhu beberapa hari saat itu sangat panas.</p> <p>Pertanyaan:</p> <ol style="list-style-type: none"> Permasalahan apa yang terjadi pada peristiwa tersebut? Mengapa hal itu bisa terjadi? Bagaimana upaya yang dapat dilakukan untuk menghidari? 	<p>4. Petugas PLN sedang memperbaiki kabel listrik yang terputus pada malam hari seperti gambar di bawah ini.</p>  <p>Memperbaiki Kabel Listrik</p> <p>Petugasnya kabel listrik lebih sering terputus pada malam hari daripada siang hari.</p> <p>Pertanyaan:</p> <ol style="list-style-type: none"> Permasalahan apa yang terjadi pada peristiwa tersebut? Mengapa hal itu bisa terjadi? Bagaimana upaya yang dapat dilakukan untuk menghidari permasalahan tersebut? Uraikan kesimpulan berdasarkan permasalahan tersebut!
3.	Fixed image	<p>3. Saat pagi hari, ayah membuat kopi. Ayah memuangkan air mendidih ke dalam gelas yang berisi kopi dan gula. Tiba-tiba gelas pecah sepele diperlihatkan pada gambar di bawah.</p>  <p>Gelas Pecah</p>	<p>5. Saat pagi hari, ayah membuat kopi. Ayah membuat kopi dengan cara mencampurkan air dan kopi di atas panci yang dipanaskan. Setelah mendidih, ayah memuangkan kopi panas ke dalam gelas. Tiba-tiba gelas ayah retak seperti gambar berikut.</p>  <p>Gelas Retak</p> <p>Pertanyaan:</p> <ol style="list-style-type: none"> Permasalahan apa yang sedang dialami ayah? Jelaskan apa penyebab gelas tersebut dapat pecah? Apa yang seharusnya dilakukan ayah agar gelas tidak pecah jika ingin membuat kopi lagi? Berilah kesimpulan atas kasus yang dialami ayah!

Quantitative assessment is carried out by ticking (√) in the "VALIDATION" column. Data obtained from the quantitative assessment were analyzed using the Content validity ratio (CVR). The results of content validity are presented in table 3.1 below.

Table 6: Instrument Validation Results

No	Score							
	V1	V2	V3	V4	V5	V6	V7	V8
1	3	4	4	4	3	4	3	4
2	4	4	4	4	4	3	4	4
3	3	4	4	3	4	4	3	4
4	4	4	3	4	4	4	4	4
5	3	3	4	4	4	4	4	3
6	4	3	4	4	3	3	3	4
7	3	4	3	4	4	4	4	4
8	4	3	4	3	3	4	4	4
9	4	4	4	4	4	4	4	3

Table 7: Data Analysis

No	S	S	S	S	S	S	S	S	S	ΣS	N (c-1)	CVI	information
.	1	2	3	4	5	6	7	8	8				
1	2	3	3	3	2	3	2	3	21	24	0.88	Valid	
2	3	3	3	3	3	2	3	3	24	24	0.96	Valid	
3	2	3	3	2	3	3	2	3	21	24	0.88	Valid	
4	3	3	2	3	3	3	3	3	23	24	0.96	Valid	
5	2	2	3	3	3	3	3	2	21	24	0.88	Valid	
6	3	2	3	3	2	2	2	3	20	24	0.83	Valid	
7	2	3	2	3	3	3	3	3	22	24	0.92	Valid	
8	3	2	3	2	2	3	3	3	21	24	0.88	Valid	
9	3	3	3	3	3	3	3	2	23	24	0.96	Valid	

DISCUSSION

The author fixes the expert's advice according to the qualitative assessment. Based on the validation carried out, the results of the instruments made were good, but more attention was paid to preparing concept indicators and question indicators. In addition, the keyword assessment guidelines are in bold to make it easier to give an assessment.

According to the validator, question number 1 is not following the indicators of problem-solving skills because it does not yet reflect a problem. In addition, the questions still use questions that are too general. Point number 4 needs to be corrected so that it is more in line with the concept of physics. The expert validator suggested confirming the story problem that incidents of disconnected cables are more common at night than during the day. Improvements to question number 5 were made to the image of broken glass in the question item to make it more contextual.

Based on table 3.2, it can be concluded that the results of a quantitative assessment of nine experts using the Content Validity Ratio (CVR) showed that the validity value was 0.88 to 0.96. It shows that the value of content validity is greater than the V value of the Aiken table, which is equal to 0.75, thus all test items are declared valid.

CONCLUSION

Based on the validation results by eight validators consisting of three expert validators and five user validators, the Thermal and Transport Concept Inventory assessment instrument has a CVR (Content Validity Ratio) value on nine items exceeding 0.75, This means that all elements have valid criteria for valid content. The CVR value obtained is between 0.88 and 0.96, so the instrument has high validity criteria. Each item and overall assessment instrument has valid criteria.

SUGGESTION

From this study, a more in-depth study can develop a STEM-based Thermal and Transport Concept Inventory assessment to measure the problem-solving abilities of high school students who have been tested for reliability and field tests.

LIMITATION

This study has limitations that only tested content validity by eight experts. It is an opportunity for the researcher to test the reliability and field test of the assessment instrument.

ACKNOWLEDGMENT

This research communication is made possible through monetary assistance from Sebelas Maret University, Indonesia.

CONFLICT OF INTEREST

The authors declare no potential conflicts of interest regarding this article's research, authorship, and/or publication.

FUNDING

Please, you write you have Funding, You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated ("This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.")

REFERENCES

- Adeleke dan Joshua. (2015). Development and Validation of Scientific Literacy Achievement Test to Assess Senior Secondary School Students' Literacy Acquisition in Physics. *Journal of Education and Practice*. 7(6): 28-42.
- Azizah,D.N., Sutopo., Zulaikah,S. (2018). Students' Problem Solving Ability through Conceptual Problem Solving Approach on Thermodynamics Material. *Journal of Education: Theory, Research, and Development*. Vol.3 No.11. 1458-1462. <http://journal.um.ac.id/index.php/jptpp/>
- Bashoori, Khoirul dan Supahar. (2018). Validity and Reliability of STEM-Based Science Literacy Performance Assessment Instruments. *Journal of Educational Research and Evaluation*. 22 (2). 222.
- Basin, K., Mađrala, M., Wąsik, M., & Małozewski, P. (2017). Interpretation of environmental tracer data for conceptual understanding of groundwater flow: an application for fractured aquifer systems in the Kłodzko Basin, Sudetes, Poland. *Isotopes in Environmental and Health Studies*. 2(1), 1–19. <https://doi.org/10.1080/10256016.2017.1330268>
- Batlolona, J. R., Baskar, C., Kurnaz, M. A., & Leasa, M. (2018). The Improvement of Problem-Solving Skills and Physics Concept Mastery on Temperature And Heat Topic. *Jurnal Pendidikan IPA Indonesia*, 7(3), 273–279.

- <https://doi.org/10.15294/jpii.v7i3.12432>
- Budiyono. (2015). Pengantar Penilaian Hasil Belajar. Surakarta: UNS Press. <http://repository.umko.ac.id/id/eprint/178/6/REFERENSI%20SELO.pdf>
- Burkholder, E., Walsh, C., & Holmes, N. G. (2020). Examination of Quantitative Methods for Analyzing Data from Concept Inventories. *Physical Review Physics Education Research*, 16(1), 10141. <https://doi.org/10.1103/PhysRevPhysEducRes.16.010141>
- Capriconia J, Mufit Fatni. (2022). Analysis of Concept Understanding and Students Attitudes towards Learning Physics in Material of Straight Motion. *Journal of Research in Science Education*. Volume 8, Issue 3, 1453-1461. <http://jppipa.unram.ac.id/index.php/jppipa/index>. DOI: 10.29303/jppipa.v8i3.1381
- Dewi, S. K., & Sudaryanto, A. (2020). Knowledge, Attitude and Behavior Questionnaire Validity and Reliability. *Proceedings of the National Seminar on Nursing*, University of Muhammadiyah Surakarta, 73-79.
- Dulari. (2015). Pengembangan Media Tutorial Berbantuan Komputer untuk Meningkatkan Keterampilan Pemecahan Masalah Fisika Siswa SMA Negeri 1 Malang. *Prosiding Pertemuan Ilmiah XXIX HFI Jateng & DIY*. <http://hfidiyjateng.or.id> accessed 5 Januari 2023.
- Eny Winaryati, Muhammad Munsarif, Mardiana, S. (2021). Circular Model of RD&D Model RD&D Pendidikan dan Sosial. Penerbit KBM Indonesia. <http://repository.unimus.ac.id/5142/1/EBook%20Circular%20Model%20RD%26D%20%28RD%26D%20Pendidikan%20dan%20Sosial%29.pdf>
- Fatih BAL, Haydeh Faraji, Y. C. (2018). Development and Reliability Analysis of Vocational Interest Inventory. *Istanbul Gelisim University Journal of Social Sciences*, 5(2), 39–53. https://www.researchgate.net/publication/329029465_Development_and_Reliability_Analysis_of_Vocational_Interest_Inventory
- Goodwin, L.D., & Leech, N.L. (2003). The Meaning of Validity in The New Standards for Educational and Psychological Testing: Implications for Measurement Courses. *Measurement and Evaluation in Counseling and Development*, 36, p: 181-191
- Hamid. (2017). Meningkatkan Kemampuan Guru dalam Menyusun Instrumen Penilaian Bentuk Tes Tertulis Melalui Bimbingan Berbasis Lembar Kerja di SMP Negeri 7 Banjarmasin. *Jurnal Ilmiah Kependidikan*, 12, 1–12.
- Husin. (2020). Peningkatan Kemampuan Guru Menyusun Instrumen Penilaian Melalui Workshop Berkala Di SMP Negeri 2 Belitang. *Jurnal Ilmiah Pro Guru*, 6(4), 461–471.
- Jati, R. P., & Aprilia, N. (2019). Evaluasi proses penyusunan instrumen penilaian kognitif. *Prosiding Symbion (Symposium on Biology Education)*, Prodi Pendidikan Biologi, FKIP, Universitas Ahmad Dahlan. <http://seminar.uad.ac.id/index.php/symbion/article/download/3542/896>
- Jua, S. K., Sarwanto, dan Sukarmin. (2018). The profile of students' problemsolving skill in physics across interest program in the secondary school. *Journal of Physics: Conference Series* 1022(1).
- Kaleka, M. U. (2016). Evaluasi kemampuan guru fisika SMA dalam menyusun instrumen penilaian hasil belajar. *Jurnal Pendidikan Matematika Dan Sains*, 4(1), 12–19. <http://journal.uny.ac.id/index.php/jpms>
- Kurniawan, B. R., Reyza, M., & Taqwa, A. (2018). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Fisika pada Materi Listrik Dinamis. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 3(11), 1451–1457.
- Kurniawati, E., Hartanto, & Zamzaili. (2017). The effect of assurance, relevance, interest, assessment, satisfaction (ARIAS) integrative and early ability model in improving the understanding of mathematical concept and troubleshooting the first middle school students in Kepahiang. *Jurnal Pendidikan Matematika Raflesia*, 2(2), 174–187
- Lavery, J. T., & Caballero, M. D. (2018). Analysis of the Most Common Concept Inventories in Physics: What are We Assessing? *Physical Review Physics Education Research*, 14(1), 10123. <https://doi.org/10.1103/PhysRevPhysEducRes.14.010123>
- Lawshe, C. H. (1975). A Quantitative Approach To Content Validity. *Personnel Psychology*, 28(4), 563– 575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Lee, M.-H, Chai, C.S., Hong, H.-Y., (2019). STEM Education in Asia Pacific: challenges and development. *Asia-Pacif. Educ. Res.* 28 (1), 1-4
- Le, L.T.B., Tran, T.T., Tran, N.H., (2021). Challenges to STEM education in Vietnamese high school contexts. *Heliyon*. <https://doi.org/10.1016/j.heliyon.2021.e08649>
- Linawati, L., Haratua T.M.S., & Hamdani. (2018). Deskripsi Miskonsepsi Siswa Pada Materi Gerak Lurus Di Sma Negeri 1 Sungai Raya. *Artikel Penelitian Fisika FKIP Untan Pontianak*, pp. 1-10. <http://dx.doi.org/10.26418/jppk.v7i11.29878>
- Lynn, M. (1986) Determination and Quantification of Content Validity Index. *Nursing Research*, 35, 382-386. <https://doi.org/10.1097/00006199-198611000-00017>
- Madsen, A., Mckagan, S. B., & Sayre, E. C. (2017). Best Practices for Administering Concept Inventories. *The Physics Teacher*, 55, 1–8. <https://doi.org/https://doi.org/10.1119/1.5011826>
- Mahirah. (2017). Evaluasi Belajar Peserta Didik. *Jurnal Idarah*. Vol 1, No 2, Hal: 258.
- Margot, K.C., Kettler, T., & Margot, K.C. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, 6, 1–16.
- Mehrnaz Mohebbi, Saaid Mahmoudian, S. A. M. (2019). Developing the Persian Version of Sensory Gating Inventory and Assessing its Validity and Reliability. *Basic and Clinical Neuro Science*, 10(6). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7253804/>
- Meiliani, M., Tanti, T., & Sulman, F. (2021). Student Resources On Newton's Law Concepts Reviewing From Gender: Identification Using Open-Ended Question. *Indonesia Journal of Science and Mathematics Education*, 324–332. <https://doi.org/10.24042/ijsme.v4i3.10177>
- Munadi. (2018) Prosiding Seminar Nasional Matematika dan Pendidikan Matematika: Integrasi Budaya, Psikologi, dan Teknologi dalam Membangun Pendidikan Karakter Melalui Matematika dan Pembelajarannya pp 559-571

- Puspitasari, R., Mufi, F., & Asrizal. (2021). Conditions of learning physics and students' understanding of the concept of motion during the covid-19 pandemic. *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/1876/1/012045>
- Rozal, E., Ananda, R., Zb, A., Fauziddin, M., & Sulman, F. (2021). The Effect of Project-Based Learning through YouTube Presentations on English Learning Outcomes in Physics. *AL-ISHLAH: Jurnal Pendidikan*, 13(3), 1924–1933. <https://doi.org/10.35445/alishlah.v13i3.1241>
- Rukman Abdullah, D. S. (2018). Validasi Metacognitive Awareness Inventory pada Pendidikan Dokter Tahap Akademik. *EJKI*, 6(1). <https://media.neliti.com/media/publications/237856-validation-of-metacognitive-awareness-in-5d041374.pdf>
- Ryu, M., Mentzer, N., Knobloch, N., (2019). Preservice teachers' experiences of STEM integration: challenges and implications for integrated STEM teacher preparation. *International Journal Technology.Des. Educ.* 29 (3), 493-512.
- Sihole, Marlon. (2015). Penerapan Model Pembelajaran Konstruktivisme Untuk Meningkatkan Kemampuan Mengidentifikasi Karakteristik Inti Atom Dan Radioaktivitas Pada Pelajaran Fisika Inti Di Kelas XII IPA 3 SMA Negeri 12 Medan T.A 2013/2014. *Jurnal Inpafi*, Vol. 3, No. 2
- Sinaga, N. A. (2016). Pengembangan tes kemampuan pemecahan masalah dan penalaran matematika siswa SMP kelas VIII. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 11(2), 169.
- Subali. (2016) Pengembangan Tes beserta Penyelidikan Validitas dan Reliabilitas secara Empiris. Yogyakarta: UNY Press. <https://unypress.uny.ac.id/buku/pengembangan-tes-beserta-penyelidikan-validitas-dan-reliabilitas-secara-empiris>
- Sukarmin., Arifin,Z., Sarwanto., Sani D, M. (2021). The Development of Concept Inventory Assessment Integrated with STEM Literacy to Measure Students' Creative Thinking Skills: A Need Analysis
- Sukenti,D., Tambak S., Charlina. (2020). Developing Indonesian Language Learning Assessment: Strengthening the personal competence and Islamic psychosocial of teachers. *International Journal of Evaluation and research In education*. Vol.9 No.4. 1079-1087. DOI: 10.11591/ijere.v9i4.20677
- Sulman, F., Sutopo, S., & Kusairi, S. (2021). FMCE-PHQ-9 Assessment with Rasch Model in Detecting Concept Understanding, Cheating, and Depression amid the Covid-19 Pandemic. *Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah*, 6(2), 297–309. <https://doi.org/10.24042/tadris.v6i2.9273>
- Sulman, F., Taqwa, M. R. A., Aminah Zb, A. Z., Rafzan, R., & Fikri, A. (2020). The Effect of Mathematical Connections on the Mastery of Probability Material. *Edumatika: Jurnal Riset Pendidikan Matematika*, 3(2), 147–157. <https://doi.org/10.32939/ejrpm.v3i2.645>
- Sutama, Sandy Ary Gilang, Fuadi Djalal. (2017). Pengelolaan penilaian autentik kurikulum 2013 mata pelajaran matematika di SMA. *Jurnal Manajemen Pendidikan*. Volume 12 No.1. 105-114
- Tong, D. H., Loc, N. P., Uyen, B. P., & Truc, C. Le. (2020). A case study of developing students' problemsolving skills through addressing real-world problems related to fractions in primary schools. *International Journal of Scientific and Technology Research*, 9(2), 2809–2818.
- wakyu dwi puspitasari, filda febrinita. (2021). Pengujian validasi isi (content validity) angket persepsi mahasiswa terhadap pembelajaran daring matakuliah matematika komputasi. *Focus Action of Research Mathematic*, 4(1), 77–90. <https://doi.org/DOI:10.30762/factor-m.v4i1.3254>
- Yuniarti Herni S, Ali Muhammad R, S Ferli Irwansyah. (2018). Verification concept of Assessment for Physics Education Student Learning Outcome, *International Journal of Engineering & Technology*. 7 (3.21) 321-332. <http://digilib.uinsgd.ac.id/id/eprint/12830>
- Zb, A., Novalian, D., Ananda, R., Habibi, M., & Sulman, F. (2021). DISTANCE LEARNING WITH STEAM APPROACHES: Is Effect On The Cognitive Domain? 6(2), 129–140.
- Zb, A., Setiawan, M. E., & Sulman, F. (2020). Pengaruh E-Learning Berbasis Schoology Berbantuan WhatsApp Group terhadap Hasil Belajar Ditengah Pandemi Covid-19. *Al-Khidmah*, 3(2), 55–60. <https://doi.org/10.29406/al-khidmah.v3i2.2282>
- Zohre Sepehri Shamloo, Seyed Sepehr Hashemian, Haniye Khoshshima, A. S. (2017). Validity and Reability of the Persian Version of the Maslach Inventory (General Survey Version) in Iranian Population. *Psychiatry Behav Sci*, 11(2). https://www.researchgate.net/publication/314070855_The_V alidity_and_Reliability_of_the_Persian_Version_of_the_Ma slach_Burnout_Inventory_General_Survey_Version_in_Irani an_Population
- Zhang, Y., & Zhang, Z. (2018). 'Kexue Wenhua' in Chinese and 'Scientific Culture', 'Science Culture', 'Culture of Science' and 'Science as Culture' in English: The Meanings and the Structure. *Cultures of Science*, 1(1), 25–37.
- Zheng, J., Huang, L., Li, S., Lajoie, S. P., Chen, Y., & Hmelo-Silver, C. E. (2021). Selfregulation and emotion matter: A case study of instructor interactions with a learning analytics dashboard. *Computers and Education*, 161, 104061. <https://doi.org/10.1016/j.compedu.2020.104061>