

DEVELOPMENT OF COMPLEX PROBLEM SOLVING TESTS IN NEWTON'S DYNAMIC FOR HIGH SCHOOL

Nurhikmah Weisdiyanti, Sahyar, Rita Juliani

Physics Education Department, Postgraduate, Universitas Negeri Medan
Jl. Willem Iskandar/Psr V, Medan, Sumatera Utara, 20221, Indonesia

*e-mail: nurhikmahweisdiyanti@gmail.com

Submitted: 11 July 2023, Revised: 17 October 2023, Accepted: 21 December 2023

Abstract. Complex problem solving (CPS) is important for students, but educators rarely provide tests that hone CPS, because the available tests are not up to standard. The aim of the study was to develop a physics test instrument based on CPS for high school that met the qualifications in the aspects of validity, reliability, discriminatory power and level of difficulty. The type of research used is the Borg & Gall research and development (R&D) model with 9 stages, namely: (1) preliminary study of problem-solving test instruments, (2) planning of instrument development and testing, (3) development of problem-solving-based test instruments, (4) initial field test by 5 material, construction and language expert validators, (5) main product revision, (6) main field test on 20 test takers, (7) operational product revision, (8) operational field test on 100 participants tests, and (9) final product revisions. This research uses a mixed method approach with qualitative and quantitative analysis. The developed test instrument consists of 15 essay test items. The results of the initial field test obtained 30 valid test items with minor revisions on the material, construction and language aspects. The results of the main field test obtained 14 test items valid and very reliable. The operational field test results obtained 14 valid and very reliable test items so that the CPS test is suitable for use and can be useful in helping educators to hone students' CPS ability.

Keywords: *Complex Problem Solving, Essay Test, Newton's Dynamics.*

INTRODUCTION

Complex problem solving (CPS) is one of the 21st century skills that needs to be provided to students according to the Global Cities Education Network Report considering the increasingly complex development of the times (Rahmawati N. R., 2019); (Situmorang & Bunawan, 2022). Students are required to be more active in honing CPS skills, where educators encourage students to develop and test their theories individually, and test theories from their peers, even when the theory is inconsistent with the actual situation, students are able to throw it away and try another theory (Situmorang & Bunawan, 2022).

The results of the Indonesian PISA study show that over the last 20 years since its release, students' reading, mathematics and science performance have not improved significantly (Pusat Penilaian Pendidikan BALITBANG, 2019). The 2018 PISA results show that the competency level of the majority of Indonesian students is below level 1 and is in 74th place out of 79 countries (Suprayitno, 2019). The research results of (Simalango, Darmawijoyo, & Aisyah, 2018) stated that students had difficulty completing the PISA test instrument in understanding questions, converting real problems and solving them.

CPS abilities can be honed through questions or tests that are linear with existing problem-solving concepts (Wardani, Arkan, & Suyudi, A, 2020), but the results of literature studies show that educators still rarely create

standard tests that measure the skills that participants must have. specific education, such as CPS abilities (Arifin & Retnawati, 2017); (Ayumniyya & Setyarsih, 2021). Educators tend to give routine questions, where the questions are more dominant in the use of formulas, even though the questions given to students are prepared at the CPS level which can trigger students to be more critical in solving problems in the questions, so that they do not experience difficulties in solving the questions. especially in terms of problem solving (Situmorang & Bunawan, 2022).

A test instrument can be said to be a good problem-solving test instrument if the test instrument is able to assess the test taker's CPS. Problem solving is assessed as an effort to find alternative solutions to a difficulty in order to realize a goal to be achieved (Durasa dan Wandung, 2021). The difference between students with low ability and high ability in solving physics problems is how they organize and use their knowledge, and connect one concept with another concept when solving problems (Wartono, Suyudi, & Batlolona, 2018).

The research "Development of Complex Problem-Solving Tests on Newtonian Dynamics Material for high School" is important to be carried out to improve complex problem-solving abilities and build students' independence in solving problems. A problem-solving based physics test instrument was developed using the Borg & Gall model (Rahmawati, Rustaman, & Dadi, 2020), and its feasibility

was analyzed in terms of validity, reliability, distinguishability and level of difficulty. The CPS test instrument developed is expected to help educators in applying test instruments to Physics learning so that students are accustomed to and skilled at solving problems based on real life problem solving.

METHODS

The type of research used is the Research and Development (R&D) Borg & Gall model. The research uses a mixed method approach. The product developed is a problem-solving based Newton dynamics material description test based on the HOTS Physics test instrument writing guide by the Ministry of Education and Culture (Wadana, 2017). The research design is shown in Figure 1.

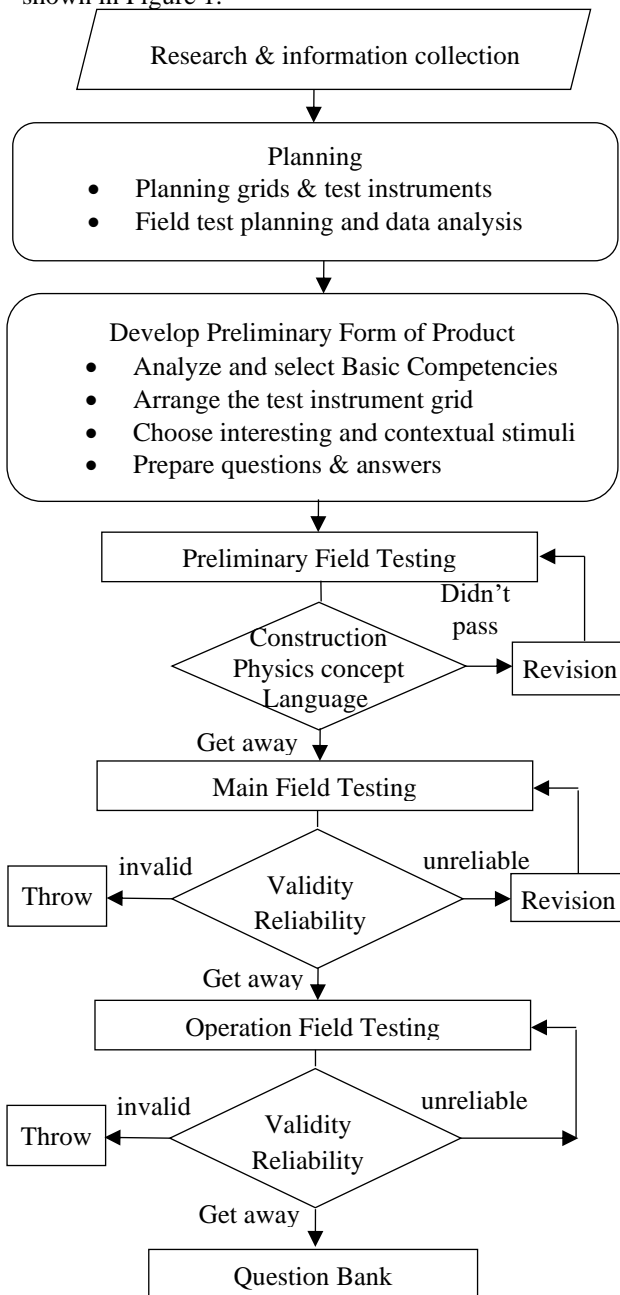


Figure 1. Design Research and development (R&D) model Borg & Gall (Gall, Gall, & Borg, 2015)

The data collection techniques used were interviews, documentation, literature studies and tests. Research data was analyzed qualitatively and quantitatively. Qualitative analysis was carried out using content validation determined by material expert agreement, test construction and language. The content validity index of test items is calculated using the validity formula according to Aiken. The characteristics of test items are analyzed quantitatively in the aspects of validity, reliability, level of difficulty and distinguishing power of test items. The validity of the calculation is obtained using the rough number product moment correlation formula with the test item which is said to be valid if $\alpha < 0.05$. For instrument reliability, Cronbach's Alpha formula is used.

CPS scoring is done by adding up the scores for each test item according to the assessment rubric used. The maximum score that tests participants can get for each indicator is 5 points. here are 5 problem solving indicators in the rubric model (Docktor, et al., 2016) so the total points per test item are 25 points. The final score scoring technique and CPS ability analysis uses equation 1 (Situmorang & Bunawan, 2022):

$$Value = \frac{(acquisition\ score)}{(maximum\ score)} \times 100 \quad \dots (1)$$

With a percentage of CPS capabilities:

$$P = \frac{x}{x_i} \times 100\% \quad \dots (2)$$

Note:

P = percentage

x = average score given by respondents per indicator

x_i = highest score per indicator

The CPS ability index of test participants is classified using the CPS ability level indicator table listed in table 1.

Table 1. CPS Capability Criteria

| Criteria | Quality Test items |
|---------------------------|--------------------|
| $85.00 \leq P \leq 100$ | Excellent |
| $70.00 \leq P \leq 84.99$ | Good |
| $55.00 \leq P \leq 69.99$ | Normal |
| $40.00 \leq P \leq 54.99$ | Not good |
| $0.00 \leq P \leq 39.99$ | Bad |

(Situmorang & Bunawan, 2022)

RESULTS AND DISCUSSION

Description of the developed Complex Problem-Solving Test

The product developed is a complex problem-solving test instrument from 15 multiple choice items to 15 description test items with Newtonian dynamics material. The test instrument is designed using stimuli that display concepts, visualizations, analogies and conclusions (Schraw & Robinson, 2011) so that it can generate complex problem-solving abilities in test participants. The test has sub questions that meet the indicators of complex problem-solving abilities. Distribution of HOTS test instruments and

cognitive level of test items that have been developed listed in table 2.

Table 2. Categories and Cognitive Levels of CPS Tests that have been developed

| CPS Level | Cognitive Level | Question Number | Percentage |
|-----------|-----------------|-----------------|------------|
| HOT (L3) | C6 | 1-4, 11, 14, | 40 % |
| | C5 | 8, 9, 13, | 20 % |
| | C4 | 6, 7, 15 | 20 % |
| MOT (L2) | C3 | 5, 10, 12 | 20 % |

The overall test instrument consists of instructions for working on the questions, question stimuli in the form of problem statements, pictures/tables, question items that meet the CPS ability indicators. The problem solving indicator was developed from Heller's problem solving indicator (Heller, Keith, & Anderson, 1992) and the problem solving indicator from the rubric (Docktor, et al., 2016), namely (a) Identification of problems and opportunities, (b) solution plan, (c) execution of the solution plan with the practical application of physics specific and mathematical procedures, and (d) evaluation and logical conclusions. The material used consists of Newton's laws and Newton's gravity. The stimulus used is technology used in everyday life in the social 5.0 era and is related to the concept of Newtonian dynamics. The instruments that have been developed are shown in Figure 2.

Sebuah *medical drone* (*Medidrone*) Multirotor diterbangkan untuk mengirim 1,6 kg suplai medis dari pelabuhan Delta Artha Bahari Nusantara (DABN) Probolinggo menuju Pulau Gilli Ketapang. Suplai medis disimpan dalam *box* medis bermassa 400 Gram berukuran 20 cm x 15 cm dan dicapit ke bagian bawah *drone* (Gambar 1.1). Berdasarkan pengamatan Badan Meteorologi, Klimatologi dan Geofisika (BMKG), kecepatan angin yang terjadi disepanjang bibir pantai Pulau Gili berkisar antar 20 km/jam sampai 40 km/jam. *Medidrone* membawa suplai medis sejauh 14 km selama 12 menit pertama dari selatan ke utara, sedangkan kecepatan angin mencapai 21 km/jam dari arah utara ke selatan. Namun, perubahan kecepatan angin yang ekstrim yakni mencapai 40 km/jam dari utara ke selatan, membuat *box* medis terlepas dari posisi seharusnya sehingga *medidrone* tidak berhasil membawa logistik medis ke lokasi tujuan. Apabila *medidrone* bermassa 1,3 kg dan membutuhkan gaya angkat minimal 6,37 N, maka:



Gambar 1.1 *Medidrone*

- Apa yang menyebabkan *box* medis terlepas dari posisi seharusnya jika ditinjau dari konsep hukum Newton?
- Berapa gaya caprit yang dibutuhkan untuk mencapit *box* medis dengan aman? (Koefisien gesekan antara *box* medis dan capitan $\mu = 0,12$)
- Bagaimana desain *Medidrone* seharusnya yang lebih aman agar *box* medis tidak mudah terlepas dari posisinya dengan perubahan kecepatan angin mendadak dari 21 km/jam hingga 40 km/jam dari arah utara? Tinjau dengan menggunakan konsep hukum Newton!

Figure 2. CPS test items that have been developed

Result

The test instrument was tested qualitatively through content validation tests and quantitatively through construction validation tests in the main field and operational fields. Content validation of test items by experts obtained an average value of 0.633 with valid

criteria. The percentage of main field test results is shown in Figure 3.

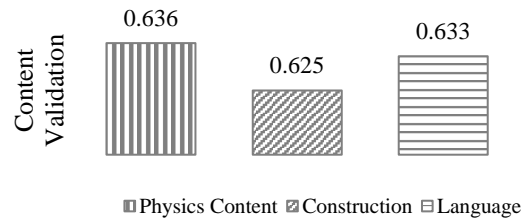


Figure 3. Test Item Content Validation Test Results

Figure 2 shows the results of the validation test of test items by experts that the test instrument is very valid in the construction aspect, valid in the content and language aspects. All test items are eligible to be tested in the main field after the material, construction and language aspects.

Validation of test instrument construction in the main field averaged 0.664 with a V index ranging from 0.421 – 0.901 at a significance level of 5%. The results of the validation test of test items in the main field class showed that 14 test items were proven to be valid and 1 test item was invalid. The percentage of main field test results is shown in Figure 4.

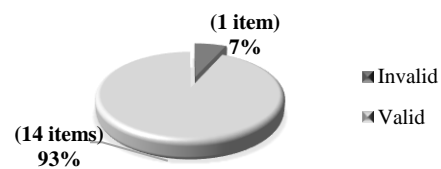


Figure 4. Percentage of Validation Test Results for Main Field Test Items

Validation of test instrument items in the operational field averaged 0.578, ranging from 0.494 – 0.838. The results of the operational field test validity analysis showed that 14 test items were tested as valid ($r_{XYcount} > 0.2108$) and there were no questions that needed to be discarded.

The reliability test results for the main field were 0.881 and for the wider field it was 0.829. The data shows that the test instrument is very reliable in the main field tests and operational field tests.

The test participants' CPS abilities are analyzed on 4 questions that have been proven to be valid, reliable and do not require a final revision process to be included in the question bank based on the results of the operational field test, namely question items number 2, 3, 13, 15. Field test participants' CPS abilities operations are tabulated in table 3.

Table 3. Test Participants' CPS Ability

| CPS Value | Percentage (%) | Criteria |
|---------------------------|----------------|-----------|
| $85.00 \leq P \leq 100$ | 0 | Excellent |
| $70.00 \leq P \leq 84.99$ | 0 | Good |
| $55.00 \leq P \leq 69.99$ | 2 | Normal |
| $40.00 \leq P \leq 54.99$ | 6 | Not good |
| $0.00 \leq P \leq 39.99$ | 92 | Bad |

Table 4.20 data shows that 92% of test participants' CPS abilities are still very low, where only 2 samples have sufficient CPS abilities and 6 others have low CPS abilities. Even though the sample chosen was from a school that has A accreditation, and the class chosen was a class with more physics abilities than students in other classes. The selection is made based on the results of discussions with the supporting teachers. The test taker's CPS abilities can be reviewed in more detail on each CPS ability indicator which is interpreted in the diagram in Figure 4.7.

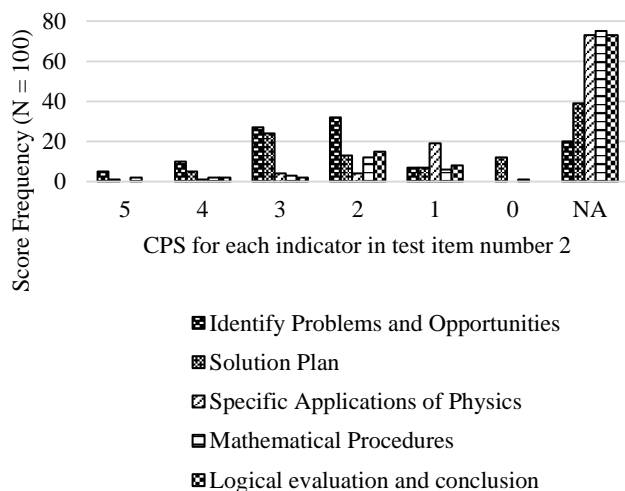


Figure 5. Problem Solving Ability for Each Indicator of Test Participants in Question Number 2

Discussion

Initial field test results show that the test instrument is valid with minor revisions in the material, construction and language aspects. Test items that do not meet the material aspects are caused by (1) the compatibility between the indicators of the test instrument, the cognitive level to be achieved with the questions and the answer choices given which are not appropriate, and (2) the stimulus in the test instrument is not yet contextual. Test items that do not fulfill the constructive aspect are caused by sentences that contain ambiguous questions and have double meaning. Test items that do not meet the language aspect are because the sentences used are not communicative (Weisdiyanti & Juliani, 2023).

The main field test results show that the CPS test instrument has been tested as valid, in the sense that the test instrument can measure the test taker's complex problem-solving abilities. Valid test items reflect that the test instrument has reliability and there is no doubt about the accuracy of the test instrument in measuring students' abilities (Sudijono, 2017). The test instrument is valid because the constructs and materials cover everything that is intended to be measured.

Validation of test instruments in the main field is higher than in the operational field. The difference is caused by the value data and answers of the main test participants being more varied than those of operational field test participants. The test instrument is more valid if the test taker's scores and answers are more varied. The results are not in line with research (Afriani, Maria, & Oktavi, 2019). The test instrument will be more valid if the number of test

takers increases. The more test takers there are, the more varied the answers they get, the more valid the instrument becomes (Afriani, Maria, & Oktavi, 2019).

The results of the main and operational field tests show that the reliability of the two test instruments ($r_{11} \geq 0,70$) is almost the same. Almost the same results show the consistency of the test instrument in measuring students' complex problem-solving abilities. It has been proven that the reliability of test instruments refers to the consistency or stability of assessment results (Reynolds, Livingston, & Willson, 2010); (Arifin & Retnawati, 2017).

Instruments that have high reliability indicate the consistency of the instrument in measuring test participants' higher-order thinking skills, the level of confidence of the evaluator in placing the test instrument as an evaluation result and the important factor in considering whether the results of the interpretation of the test instrument can be operationalized (Sukardi, 2008). A reliable instrument will obtain results that are not much different when used in other schools (Marwan, Khaeruddin, & Amin, 2020). The consistency of the test instrument refers to the precision of the test taker's scores and answers.

The results of the three field tests showed that there was an increase in the quality of the question items after being revised 3 times, so that of the 15 test items developed, there were 14 test items that were suitable for inclusion in the question bank and used as problem-solving based test instruments for SMA/MA level. The research is in line with (Kurniawan & Taqwa, 2018) which shows that 7 test items representing four indicators of dynamic electrical material have good validity and reliability.

The CPS ability indicator for identifying problems and opportunities is the indicator that most test takers answer, especially at score 2 at the cognitive level of Application (C3), Analysis (C4), Evaluation (C5) and creating (C6). Further analysis shows that most test takers identify problems and opportunities by creating known variables and are asked about them in the questions and rarely sketch pictures of objects along with diagrams of the forces acting on the object, so that the identification of the problem given is incorrect or is unrelated. with physics concepts. Problems are caused because test takers are used to identifying problems in conventional ways (Docktor, Strand, Mestre, & Ross, 2015).

Indicators of solution plans with a physics approach include indicators that are difficult after execution of the solution plan, evaluation and conclusion. Test takers tend to still have difficulty making plans to solve questions. The planning stage requires students to understand the concept as a prerequisite for solving the problem, because in making a plan to solve a problem, students must be able to connect one concept with another (Farikh & Haryani, 2022). The low level of achievement in mastering aspects of planning solutions using a physics approach confirms that students are often only oriented towards memorizing formulas and do not have the ability to choose formulas that suit the context of the problem (Aristiawan, 2022).

The ability to execute a solution plan by applying specific physics and mathematical procedures is related to the test taker's ability to identify problems and opportunities and plan solutions. Test takers who correctly identify problems and opportunities, especially those who are able

to visualize problems and the forces acting on objects, can execute solution plans correctly too. On the other hand, if you make a mistake in describing the forces acting on an object, there will be an error in formulating the problem according to a physics approach. Test participants who are able to plan solutions using a physics approach, complete with analysis of the forces at work, will be correct in applying the specific formulas that apply.

Evaluation and logical conclusions are the lowest CPS for test participants. Several indicators indicate low mastery of aspects of evaluating solutions, namely students do not make diagrams or sketches that describe questions, students do not carry out checks after completing mathematical operations, students do not come up with or look for arguments that support the answers students get (Docktor, et al., 2016). The majority of test takers directly enter the existing numbers into formulas that they have already memorized, while most of the test instruments tested require a combination of several physics' concepts such as a combination of Newton's laws and uniformly changing circular motion. The low achievement of evaluation and logical conclusions confirms that students are often only oriented towards memorizing formulas and then immediately calculating existing data without any process of selecting related concepts so that students do not have the ability to choose formulas that are appropriate to the context of the problem (Aristiawan, 2022).

Overall data shows that more than 60% of test takers chose not to answer all test items so as to get an NA score. Test participants choosing not to answer could be due to lack of time in answering the test instruments and the test participants' low KPM, but if we look at the test participants who successfully answered almost all of the test instruments, the main cause is more about the students' low KPM in physics. The low CPS ability of test participants, even though the samples taken were from high schools that have A accreditation and classes that are considered to have higher physics abilities in other classes, could be due to test participants not being used to solving problem-solving-based questions. Increasing problem solving abilities can be done by creating appropriate learning to facilitate students to improve their ability to solve problems. Improvements to test instruments are also needed by carrying out analysis both qualitatively and quantitatively to obtain quality questions.

CONCLUSION

The complex problem-solving test instrument on Newtonian Dynamics for High School was developed in the form of a 15-item essay. The conclusion of the analysis and discussion results is that the test instrument is suitable for use as a tool for measuring students' complex problem-solving abilities with valid characteristics according to material, construction and language experts with an average of 0.633 in the valid category and has obtained empirical evidence through construction validation with 14 valid items in the main field test and in the operational field test. The test instrument was also proven to be reliable with a value of 0.881 ($r \geq 0.70$.) for the main field test results and 0.829 ($r \geq 0.70$) for the operational field test results, so that the CPS test instruments can be implemented by

educators in the teaching and learning process so that later it can improve students' CPS abilities.

REFERENCES

- Afriani, E., Maria, H. T., & Oktavi, E. (2019). Pengembangan Tes Higher Order Thinking Skills (HOTS) Materi Gerak Lurus Berubah Beraturan untuk SMA. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, 8(3), 1-12.
- Arifin, Z., & Retnawati, H. (2017). Pengembangan Instrumen Pengukur Higher Order Thinking Skills Matematika Siswa SMA Kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 12 (1), 98-108.
- Aristiawan. (2022). Profil Kemampuan Pemecahan Masalah Fisika Siswa SMA Menggunakan Soal Essay. *Jurnal Tadris IPA Indonesia*, 2(1), 45-55.
- Ayumniyya, L., & Setyarsih, W. S. (2021). Profil Kemampuan Berpikir Tingkat Tinggi Siswa SMA dalam Pemecahan Masalah Pada Materi Hukum Newton. *Inovasi Pendidikan Fisika*, 10(1), 50-58.
- Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., . . . Yang, J. (2016). Assessing Student Written Problem Solutions: a Problem-Solving Rubric with Application to Introductory Physics. *Physical Review Physics*, 12, 1-18.
doi:<http://dx.doi.org/10.1103/PhysRevPhysEducRes.12.010130>
- Docktor, J. L., Strand, N. E., Mestre, J. P., & Ross, B. H. (2015). Conceptual problem solving in high school physics. *Physics Education Research*, 11, 1-13. Retrieved from <http://dx.doi.org/10.1103/PhysRevSTPER.11.020106>
- Farikh, N. I., & Haryani, S. (2022). Pengembangan Instrumen Tes Berbasis Pemecahan Masalah Pada Materi Redoks. *Chemistry in Education*, 11(1), 79-85.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2015). *Applying educational research: How to read, do, and use research to solve problems of practice 7th Edition*. New York: Longman.
- Heller, P., Keith, R., & Anderson, S. (1992). Teaching problem solving through cooperative grouping. individual problem solving. *American Association of Physics Teachers*, 60(7), 627-636. doi:<http://dx.doi.org/10.1119/1.17117>
- Kurniawan, B. R., & Taqwa, M. R. (2018). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Fisika pada Materi Listrik Dinamis. *Jurnal Pendidikan Teori, Penelitian dan Pengembangan*, 3(11), 1451-1457.
- Marwan, M., Khaeruddin, & Amin, B. D. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Pada Bidang Studi Fisika. 2, 116 - 119. Retrieved from <http://ojs.unm.ac.id/semnasfisika>
- Pusat Penilaian Pendidikan BALITBANG. (2019). *Pendidikan di Indonesia: Belajar Dari Hasil PISA (Programme for International Student Assessment*

2018. Jakarta: PUSPENDIK BALITBANG KEMENDIKBUD.
- Rahmawati, N. R. (2019). Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah pada Materi Kelistrikan dan Kemagnetan. *Seminar Kontribusi Fisika*, 94-103. doi:978 - 6 0 2 - 6 1 0 4 5 - 6 - 4
- Rahmawati, Rustaman, N., & Dadi, R. (2020). Pengembangan Instrumen Tes KPM pada Materi Kelistrikan dan Kemagnetan. *Seminar Kontribusi Fisika*, 96-103.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2010). *Measurement and assessment in education (2nd Edition)*. New Jersey: Pearson.
- Schraw, G., & Robinson, D. R. (2011). *Assessment of higher order thinking skills*. Charlotte: Information Age Publishing. Retrieved from <https://eric.ed.gov/?id=ED527631>
- Simalango, M. M., Darmawijoyo, & Aisyah, N. (2018). Kesulitan Siswa dalam Menyelesaikan Soal-Soal Pisa Pada Konten Change And Relationship Level 4, 5, dan 6 di SMP N 1 Indralaya. *Journal Pendidikan Matematika*, 12, 43-58. doi:2549-1040
- Situmorang, H., & Bunawan, W. (2022). Pengembangan Instrumen Tes Untuk Mengukur KPM Siswa Pada Materi Gerak Parabola. *Jurnal Inovasi Pembelajaran Fisika*. *Jurnal Inovasi Pembelajaran Fisika*, 10 (3), 44-53. Retrieved from <http://jurnal.unimed.ac.id/2012/index.php/inpafi>
- Sudijono, A. (2017). *Pengantar Evaluasi Pendidikan*. Jakarta: PT. RajaGrafindo Persada.
- Sukardi. (2008). *Evaluasi Pendidikan Prinsip dan Operasional*. Jakarta: Bumi Aksara.
- Suprayitno, T. (2019). *Pendidikan di Indonesia: Belajar dari Hasil PISA 2018*. Jakarta: Pusat Penilaian Pendidikan BALITBANG KEMENDIKBUD.
- Wadana, I. W. (2017). *Modul Penyusunan Instrumen tes Higher Order Thinking Skills (HOTS)*. Jakarta: Direktorat Pembinaan SMA.
- Wardani, A., Arkan, M., & Suyudi, A. (2020). Pengembangan Instrumen Tes KPM Pokok Bahasan Listrik Dinamis. *Jurnal Kependidikan Betara*, 1(1), 14–19.
- Wartono, Suyudi, A., & Batlolona, J. R. (2018). Students' Problem Solving Skills of Physics on the Gas Kinetic. *Journal of Education and Learning*, 12(2), 319-324. doi:10.11591/edulearn.v12i2.8424
- Weisdiyanti, N., & Juliani, R. (2023). Development of Critical Thinking Skills-Based Physic Test Instruments for High School. *ICosTA*, 1-14. doi:<http://dx.doi.org/10.4108/eai.1-11-2022.2326211>