



## Review of problem-solving measurement: an assessment developed in the Indonesian context

Azizul Ghofar Candra Wicaksono\*

*Doctoral school of education, University of Szeged, Szeged, Hungary*

*ORCID: 0000-0002-7094-5208*

Erzsébet Korom

*Doctoral school of education, University of Szeged, Szeged, Hungary*

*ORCID:0000-0001-9534-8146*

---

<b>Article history</b>	<p>The accuracy of learning results relies on the evaluation and assessment. The learning goals, including problem solving ability must be aligned with the valid standardized measurement tools. The study on exploring the nature of problem-solving, framework, and assessment in the Indonesian context will make contributions to problem solving assessment in Indonesian educational learning system. This review involved 32 studies that focus on problem-solving test development conducted in Indonesia and have the Indonesian version of the test. All tests are in the scope of certain subjects (mathematics, science, physics, and chemistry) and administered grade 7 to undergraduate level. Each test revealed a good value of reliability. Most of them have acceptable reliability score (r-value between .60 and .80) and high-reliability score (<math>r &gt; .80</math>). Besides, they also showed content and construct validity (acceptable r value in Pearson product moment analysis and INFIT MNSQ index), but additional analysis is needed to fully develop the tests' empirical evidence. All the tests are categorized as domain specific problem solving which focus on mathematics, science for junior high school, physics, chemistry, and biology. In addition, the topic coverage in the test should be improved and further studies about the measurement of problem-solving and test development are needed in the case of the Indonesian context.</p>
<b>Received:</b> 25.10.2020	
<b>Received in revised form:</b> 11.07.2021	
<b>Accepted:</b> 17.07.2021	
<b>Key words:</b> Word; Problem-solving; Assessment; Indonesian context	

---

### Introduction

Based on the 21st-century framework, education is directed to learning and innovation skills such as problem-solving (Partnership for 21st-century skills, 2009). As a consequence, the educational practice is trying to implement a learning strategy based on the stated problems (Ferreira & Trudel, 2012; Hung et al., 2012) and use assessment for evaluating problem-solving skills. In recent days, there are many types of problem-solving assessment tools. Since the theory of problem-solving develops rapidly through many research and

---

\* Correspondency: [azizul.wicaksono@edu.u-szeged.hu](mailto:azizul.wicaksono@edu.u-szeged.hu)

studies, the assessment tools of problem-solving also varied. There is various problem-solving assessment in terms of their forms (i.e., computer and paper-based test), settings, and background frameworks. However, the varying study of problem-solving led to some confusion (Greiff, 2012). Different theories sometimes explain different purposes and views in defining problem-solving. Indeed, exploration of the terminology of problem-solving must be made to get the necessary knowledge for assessment study. One theory may contradict with others, which makes the identification process becomes critical for starting problem-solving assessment study.

There are many kinds of well-known problem-solving tests available for educational purposes. Starting with complex problem solving (CPS) that was developed in different assessment tools like Genetclab and MicroDYN (Sonnleitner et al., 2012). It emphasizes the ability in making connections with previously dynamic and unknown system. The complex problem solving means that if the problem situation changes, successful integration and exploration of information or knowledge are gained by the user intervention or environment regularities. Programme for International Student Assessment (PISA) also used CPS in assessing educational system among countries, but they focus on the interaction between problem and problem solver (OECD, 2014) and now they are focusing on a collaborative term of individual to be engaged in the team for solving the problems. Therefore, students are predicted to be able to establish an effective team organization while solving certain problems (OECD, 2017). The other term of problem-solving is firstly raised by Polya that specifically addressed problem-solving in mathematics. Then, many studies developed a problem-solving framework and assessment in a specific domain and revealed that problem-solving has a strong relation with knowledge (Dermitzaki et al., 2009; Liao, 2002). The rationale of domain-specific problem solving is that some problems that happened in a specific situation can be solved only by experts who have a strong background in that field. Moreover, in teaching practice, many educators use specific-domain problem-solving in teaching problem solving based on their subjects (Gok, 2010). As there is inconsistency in the problem-solving terminology, then further problem solving impedes in general term and also narrows in domain-specific areas such as science, mathematics, management, and technology (Sugrue, 2005).

The evaluation of problem-solving skills had been done in many countries as PISA launched cognitive and collaborative problem-solving assessment. However, some countries did not participate in PISA problem solving including Indonesia. The profile of students regarding problem-solving in Indonesia is unknown yet since there are no comprehensive studies in the problem-solving survey. Despite many studies conducted in problem-solving based learning (Asyari et al., 2016; Iswandari et al., 2017) to improve student's problem-solving skills in Indonesia, there are few studies focused on problem-solving assessment in Indonesian context. Thusly the evaluation and review of student's problem-solving assessments are needed to be prioritized to get a deeper understanding and develop further recommendations for developing the problem-solving assessment.

## **Theoretical Background**

### ***Definition of problem-solving: the general term of 'problem' and 'problem-solution'***

To get a deep understanding of problem-solving, many philosophies and psychologists think back about the root definition of problems and how to solve them. In the present study, based on the Cambridge dictionary, the word 'problem' is described as a harmful or

unwelcome situation or matter needing to be dealt with. It is also defined as a person, situation, or thing that needs attention and needs to be overcome. The term 'problem' conjugates with the word 'solve' which mathematicians refer to as problem-solving (Schoenfeld, 1987). The others argue that the problem is related to a condition, they mentioned that the problem was a situation in reaching some goals (Glaser et al., 2009). The solution can be a goal for problems and every problem solver is seeking it. Besides, it is important to emphasize that the problem has more conceptual depth than just a question. In the 'problem' standpoint, there should be a clash between belief and claims, fact and thought, or between people's thoughts (Carlson & Bloom, 2005).

There are two classes of problems, firstly it was a well-defined problem that its goals, in which the way to solve the problem and the obstacle in achieving the solution are well known, based on the knowledge and information given. Secondly, ill-defined problems that are characterized by the lack of path solutions. There is no exact solution, so these problems can be solved in many ways and the task to solve the problems becomes more challenging (Davidson, Sternberg, & Sternberg, 2003). Multiple arguments and problem representation may be the best present for ill-defined problems to find the right solution. Moreover, the problem is only a problem if we do not know how to deal with it. Solving the problem then becomes critical even done in simple ways, comfortably by routine or familiar procedures or when it requires complex conditions. Thus, solving the problem requires high mental activities in case they must be smartly identified and the best solution to be specified. The act of solving the problem then involves a mental and cognitive process. Even many neurologists propose a constructive model of brain function related to problem-solving process. It arises from the interaction sub-network and system level of brain that coordinate together in multifaceted cognitive process (Bartley et al., 2018).

Solving the problem is accounted for as variant formulations of seeking the truth and building a foundation of knowledge. It is a principal unit of achievement (Nickles, 1988). However, a problem solver itself imposes an inquiry process in figuring out the truth and finding the formulation of good problems, searching solutions to it, and testing those solutions. Those principles are also similar to what Polya (1945) explained; problem-solving involved some activities including understanding the problem, developing a plan, carrying out the plan, and looking it back (Schoenfeld, 1987). He was describing problem-solving activities as a linear progression from one stage to the next level and advocate that was the way of solving the problem. Some studies describe problem-solving behavior in different phases of metacognitive activities such as orientation, organization, execution, and verification. The problem solver can shift to the next phase of the solution when metacognitive decisions result in real behavior or cognitive action (Carlson & Bloom, 2005). Moreover, the stages in solving the problem in which every problem solver must be able to recognize or identify the problem, mentally define the problem, develop a strategy for the solution, organize the knowledge about the problem, namely, using both mental and physical activities for solving the problems, monitoring their progress, and evaluating their solution for accuracy (Davidson, Sternberg, & Sternberg, 2003). Furthermore, he said that those stages are not always processed sequentially through all stages. The successful problem solvers are indeed quite flexible. Sometimes the solution to a problem gave rise to another problem. And again it needs to be solved through the problem-solving cycle.

### ***Intelligence domain-general problem solving***

Domain general problem solving reflects a necessary skill for adapting the cross-cultural and complex problem environment in our present society. It connects on cognitive and metacognitive process such as reasoning, making planning and decision, information processing, meta-strategic thinking, and evaluation of knowledge (Greiff et al., 2014). Domain-general problem-solving in educational contexts is relevant to skills in the daily life of 21st century. For example, in the secretary job, organizing and scheduling a business meeting relies on specific skills. In order to fully accomplish the task, the combination of knowledge, strategies and experience will usually suffice. But, when the new situation arises or something unexpected happened, the domain general problem solving comes into play. It has a purpose to adapt to the new situation, explore new solution, and make correct decisions and fast adjustments (Greiff & Neubert, 2014).

### ***Complex problem solving (CPS)***

Complex problem solving is one of the domain-general issues that is characterized by the successful interaction with a dynamic task environment and gained through integrated information in that process (Wüstenberg et al., 2012). It is also described as successful interaction with non-routine and dynamic changing of the environment. It represents a variety of situations that happen in daily life (Rudolph et al., 2017). CPS emphasizes a complex cognitive system, like planning the action, developing a strategy, acquiring the knowledge, and making the evaluation that led to specific goals (Funke, 2010). The basic knowledge is needed to identify the most relevant structure of the problem and assists in covering possible states of the problems, as well as the problem structures and schemas. The benefit of knowledge is in connection with fast prediction and problem analysis, which makes problem solver be able to accept, reject, or modify previous assumptions (Greiff, Fischer, et al., 2015). However, the main point in complex problem solving is not all of the information is necessary to solve the problem. It is more pertinent to a process in generating information processes with adequate strategies and procedural abilities to control the given system (Wüstenberg et al., 2012).

As part of domain-general problem solving, CPS is independent of the person's prior knowledge. The knowledge is important, but in the CPS task most information is not needed because it will result in decreasing the process of controlling the system and integrating knowledge (Greiff, Fischer, et al., 2015). The requirement of problem solver in complex problem-solving scope, includes (1) the complexity related to reducing information, (2) transparency in generating information, (3) interconnectedness in building problem's model, (4) dynamics in forecasting and controlling future development, and (5) polytely, reaching more than one goal in complex situation (Funke, 2010). CPS is comprised of two phases, knowledge acquisition and knowledge application (Dindar, 2018; Funke, 2010; Greiff et al., 2014). In the knowledge acquisition phases, the problem solver identifies the dynamic and variables provided by the system and try to develop a representation of the stated problem. They need to explore and understand the complex system. In the knowledge application phase, the test takers transform the complex system into specific state and control it by updating their knowledge.

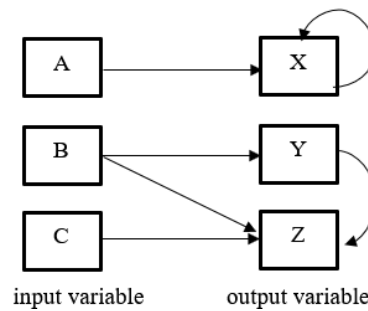
The assessment of CPS is varied among studies. In some cases, their results reported many variations and differences. Based on a meta-study conducted (Stadler et al., 2015), the measure of CPS is coded into three different measurements involving classical CPS measurement, single complex system (SCS), and multiple complex systems (MCS). A

classical measurement of problem-solving, for example, microworlds, emulated real-world problems (Greiff, Fischer, et al., 2015). At the very beginning, microworlds present multiple problem situations very broadly related to society such as governing a small town. Then it is replaced by specific problem situations in the new context “microworlds taylorshop” that represent problems in the retail business. All of the variables included in the taylorshop are very similar to real-world problems, even if the test taker manipulates those variables, it can give results in the microworld similar to the real condition (Greiff, Stadler, et al., 2015). Hence, the classical problem-solving test may be simple but necessary to give a simulation of realistic problem situations. There are some limitations to the use of classical measurement in assessing problem-solving. It has resemblance to real world more and did not have systematic theoretical framework in their problem-solving test construct (Funke, 2001). Another weakness in classical measurement is that one task comprises several interrelated items. The problem solution is influenced by many other variables and participant’s previous action. Thus, the items are hardly independent of each other.

Besides the classical measurement, problem-solving tests also developed in the form of a single complex system (SCS). One of the popular SCS problem-solving tests is multifluxplex firstly designed by (Kroner et al., 2005) based on the linear structural equation system. This test is considered as a one-item-test because the scenarios are generally constructed by one specific system configuration. Indeed, every indicator assessing every domain in this test during the system exploration is related to the same structure (Wüstenberg et al., 2012). Even when the test taker does the different tasks of test series with different goals, it still depends on the same system structure. In this program, participants will explore some tasks that have some additional effects in generating knowledge. The multiflux have four principles in their system (Christ et al., 2020; Wüstenberg et al., 2012), they are rule identification strategy, causal knowledge, rule knowledge, and rule application.

The third type of CPS is multiple complex system (MCS) that has some differences compared to classical and SCS in the term of the variables. In contrast with classical and SCS problem solving, MSC uses multiple and independent items to assess problem-solving ability (Wüstenberg et al., 2012). One of the MCS problem-solving tests widely used in the assessment process is microDYN. This test is a computer-based assessment with multiple independent items and multiple control roles in each item (Greiff, 2012). This test consists of 8-10 complex items with 3 different input variables (denoted as A, B, C) and three output variables (denoted X, Y, and Z) (Rudolph et al., 2017). The test takers make interconnections between input variable and output variable by manipulating it. The test is modified with the pattern when input variable influences output variable, or when output variables? influence each other. As is seen in figure 1, for the example, the input variable ‘A’ can influence output variable ‘X’ while variable ‘B’ influences variable Y or Z. Even in the same output variable, variable ‘Y’ can influence variable ‘Z’. Then variable C influences variable Z and so on. This complex interrelation between variables is worked based on the specific equation in which all the possible relations are equal to the number of the output variables (Wüstenberg et al., 2012).





**Figure 1.** The scheme of variable involved in MicroDYN test (Wüstenberg et al., 2012)

### *PISA test of problem-solving: creative and collaborative problem solving*

In the domain-general problem solving, one of the most well-known assessment is creative and collaborative problem solving conducted by the Programme for International Student Assessment (PISA). They made assessment tools to compare the educational system worldwide (OECD, 2016). In 2012, PISA introduces cognitive problem-solving assessment framework with a computer-based test design in which the students are faced with daily routine real-life problems such as using a new mobile phone, how to fix lamp and electricity problems, or finding some locations with different paths/routes (OECD, 2014). However, sometimes they involved tasks that are non-routine for the students or test takers but still, the problems are pointed to general knowledge strategy. The framework of PISA creative problem solving is presented in three distinct aspects: the nature of problem situations, problem-solving process, and problem context. The PISA creative problem-solving test comprised of static and interactive tasks. The static task is mainly focused on decision-making problem tasks with a different type of static units. All of the units are delivered on a computer video game mechanics (Dindar, 2018). The interactive unit in the PISA test belongs to complex problem solving, microDYN, and finite-state automata (OECD, 2014). They use ‘control’ and ‘exploration’ of an unknown system for student problem-solving tasks. Four units in the test are microDYN units and six of them are finite-state automata. The measurement of PISA’s creative problem-solving test is classified into seven proficiency levels (below 1, 1, 2, 3, 4, 5, 6, and 7) based on the items or units the test taker can solve.

Later in 2015, PISA announced the different types of problem-solving tests, and came the new term: collaborative problem-solving. This construct has a rationale based on workplace demand in which problem-solving is not needed for an individual task only, but a whole team project. In specific conditions where no individuals are able to solve the problems, collaboration and teamwork become essential to find the solution and reach the goal. The collaboration will combine the ideas, methods, and efforts in response to the problems (Care et al., 2016). Even in the school environment, collaborative interaction between students in class performance will result in better achievement and enhance their ability of solving tasks (Fawcett & Garton, 2005).

The collaborative problem solving is described as (OECD, 2017, p. 47): “*the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing understanding and effort required to come to a solution and pooling their knowledge, skills, and efforts to reach that solution*”. In collaborative problem solving, PISA adds three collaborative competencies while keeping using four problem-solving processes from a creative problem-solving framework. The three competencies are (1) establishing and maintaining shared understanding, in which the test takers try to identify knowledge and the

group's members' perspectives about the problem, (2) taking appropriate actions to solve the problem, in these steps, the test takers identify the activities that can be done in the team to solve the problems and achieve the solution, and (3) establishing and maintaining team organization. Here, the test takers have to understand the member and the agent's role in monitoring the activity and facilitating the changes to reach a better performance in solving the problems (OECD, 2017).

In PISA collaborative problem-solving test, they have started to include a new aspect of collaborative activities, with (1) hidden profile tasks or a jigsaw. In this task, the test taker will be put in a group (with a computer-based agent) and gets a task in which each group member has different information and skills. They need to collect and use the knowledge or information to solve the problem together. In the test system, the test taker will be forced to depend on one another to arrive at the solution, and hence collaboration between members is required. (2) the consensus-building task, in this, all of the group members contribute to giving opinions and making arguments toward the problems, the decision must be taken after considering the views of all group members, even when an argument is not fully altered and another argument seems dominant, they all lead to a group solution, and (3) negotiation task, where not all members share the same idea, and they need to negotiate which ideas can be selected as a final solution that satisfy individual members and the whole group (OECD, 2017).

### ***Knowledge based domain-specific problem solving***

The problem-solving complexity is distinctive not only for its characteristics, but also its application in a certain condition or problem situation. Despite being manifested in intelligence and general domain ability, some studies put problem-solving in a special, domain-specific category, based on its context. This term appeared a decade ago when Polya (1945) used problem-solving in mathematics education. He uses problem-solving in stating mathematics learning, and later he describes problem-solving in several steps, starting with the understanding of the problem, devising a plan, carrying out the plan, and evaluating the problem solution. Then, many educational researchers used problem-solving in a specific subject in various learning situations (Mukhopadhyay, 2013; Yu et al., 2010). Acquiring problem-solving for students (i.e., senior high school) is provided by teachers through the teaching-learning processes pertaining to different subjects. They mainly deliver problem-solving in domain specificity of situation specificity based on the learning topic.

The main feature of domain-specific problem-solving lies in the position of knowledge for problem-solving construction. Hence, back to the core of the main function of problem-solving in giving a solution to a problem, the knowledge and strong information related to the problem are needed. The solution depends on the information processed (Walker et al., 2016). Wolff, 2017 expresses the strong effect of knowledge based on solving problems. This knowledge is not specified only for the content information, but also acts in organizing and representing information retrieval to facilitate an efficient problem-solving process. The other study argued that problem-solving is one of the human competencies that are considered domain-specific, it has a relatively narrow domain (Sternberg, 2018). As a competency, most people had mastered it in a specific domain and less in others. Furthermore, some studies focused on joint action in solving a problem, and their specific condition with the problem exists suggesting the role of a situated condition in problem-solving. That makes an argument if someone is able to solve one problem, they cannot guarantee to solve another problem in a different situation. Thus, the knowledge-based problem-solving can be broken down into two

categories, that is, what information in subject-related problem solving of an individual possesses and how they use or process it. The main core of this logic is the difference between knowledge attributes (facts, theories, principles, definitions, and strategies) plays a different role in the way to solve the problem starting from exploring the context, discovering information, building hypothesis, and confirming or verifying the solution (Csapó & Funke, 2017).

In the framework of domain-specific problem solving, Rausch & Wuttke(2016) documented phases of a problem-solving approach. They identify four main steps including (1) identifying information gaps and needs for action, (2) processing information when the information and knowledge related problem are stored, interpreted, and used for understanding the problem and making a decision (Lachman et al., 2015), (3) arriving at a well-establish solution. Based on the available information and cognitive processing, a solution can be proposed with strong analytical calculation and (4) communicating decision, despite taking actions for solving the problem, one step that is needed to do is communicating the result of the problem solution. Hence communicating a solution in oral or written form and making anyone aware of it becomes an important facet of domain-specific problem-solving Schoenfeld(2013) explains if a person wants to engage in a goal-oriented activity such as problem-solving, he or she needs to make a series of activities such as stating the goals, maintaining the individual knowledge or resources of his or her disposal, developing individual beliefs and orientation, and making decisions.

Domain-specific problem solving has a different position in real-life problem-solving activities. While domain-general plays an important role in the daily routine real-world context as basic intelligence skills, domain-specific problem-solving is mostly used in educational training for the teaching of problem-solving. In a complex and specific problem situation, the procedure for finding solutions and making decisions are the same with the specific domain problem-solving approach. That makes a sense, in the real world, a complex problem should be given provided for the experts for them to solve it. Indeed, they also use a general domain framework to solve the problem efficiently. In conclusion, both general and specific-problem solving are needed in solving a real-world problem when domain-general plays as basic intelligence in individual skills and specific-domain supports through providing a problem-solving approach with comprehensive knowledge and information.

## **Method**

### ***Data Source***

The research included in this review is restricted to problem-solving test development research. All of the studies were achieved from a comprehensive search through databases DOAJ, Research Gate, ERIC, and Google Scholar containing studies published from January 2010 to June 2020. The search strategy is varied among databases, but it commonly includes keywords such as “assessment”, “test”, “problem-solving”, “mathematics and science”, “validation/validity”, or “Indonesia”. Systematical search was conducted in entering the combination of keywords in the databases both in English and Indonesian. All of the studies gathered were conducted in Indonesia and administered the Indonesian version of problem-solving tests. All published studies in journals and conference proceedings are involved in this review.



### ***Inclusion and exclusion criteria***

The studies included in this review should meet the following criteria: (1) the tests are standardized (2) the sample number of the study was reported (3) empirical results in validity and reliability were declared (4) the tests were projected to high school and university students. The total amount of 93 studies about the Indonesian problem-solving test was found. 61 (65.59%) of them were excluded because the study only showed test design without any report on the empirical study. In the end, a total amount of 32 (34.41%) studies were included with criteria in showing test structure, validity, and reliability results.

### ***Data coding and analysis***

The features related to the focus of the study were coded including (a) the topic/content of the test focus (e.g., mathematics, science, physics, chemistry, or biology), (b) the test development quality that indicated by validity and reliability value, (c) grade distribution, and (d) test framework, indicated by the theoretical background in the developed test items. Furthermore, all of the studies were analyzed descriptively.

### **Result and Discussion**

Out of the 32 published studies used in this review, 1 study (3.13%) was found in 2012, followed by 2 (6.25%) studies in 2014, 1 (3.13%) in 2015, 4 (12.50%) in 2016, 7 (21.87%) in 2017, 9 (25.12%) in 2018, 5 (15.63%) in 2019, and 3 (9.37%) in 2020. Tests used a specific subject and a different problem-solving framework. Each problem-solving framework is translated into questions in multiple-choice and essay form. The total items developed in the studies are varied from 5 to 103 items. All tests are in the paper-based form. The details of the problem-solving test founded in Indonesia are described in Table 1.

Table 1. The developed problem-solving test in Indonesia

No.	Author (s)	Year	Subject	Test items		Number of participants	Reliability	Type of validity	
				N	type				
1	Sinaga	2016	mathematics	5	essay	90	.650	Content	&
2	Pardimin et al.	2017	mathematics	5	essay	90	.803	Content	
3	Anggraeni et al.	2018	mathematics	15	essay	40	.732	Content	
4	Wahyuningrum	2014	mathematics	5	essay	122	.595*	Content	&
5	Zulkarnain et al.	2018	mathematics	5	essay	30	.690	Content	&
6	Ariawan	2016	mathematics	6	essay	78	.640	Content	&
7	Bidasari	2017	mathematics	12	essay	36	.871	Content	&
8	Novita	2012	mathematics	15	essay	25	.737	Content	
9	Februarini et al.	2017	mathematics	6	essay	38	.800	Content	&
10	Putra	2017	mathematics	5	essay	10	.580*	Content	&
11	Subekti et al.	2014	mathematics	5	essay	29	.761	Content	&
12	Hendriyana et al.	2018	mathematics	5	essay	66	.884	Content	&
13	Peranginangin et al.	2019	mathematics	4	essay	32	.751	Content	

14	Irawati et al.	2018	mathematics	4	essay	5	.780	Content & Construct
15	Alfika et al.	2018	science	5	essay	10	.697	Content
16	Hidayat et al.	2017	science	15	essay	35	.880	Content & Construct
17	Lutfi et al.	2019	science	15	essay	30	.835	Content & Construct
18	Pratiwi et al.	2015	physics	15	essay	35	.917	Content
19	Nadapdap et al.	2017	physics	52	multiple choices	281	.800	Content & Construct
20	Syifaulyah	2019	physics	10	essay	30	.880	Content
21	Lestari et al.	2019	physics	5	essay	40	.839	Content & Construct
22	Sinensis et al.	2019	physics	20	essay	52	.759	Content & Construct
23	Kurniawan et al.	2018	physics	8	essay	60	.670	Content & Construct
24	Yulianti et al.	2018	physics	103	Multiple choices	34	.780	Content & Construct
25	Sutiadi et al.	2016	physics	24	essay	37	.431*	Content & Construct
26	Rifa'i et al.	2018	physics	5	essay	101	.621	Content
27	Savitri et al.	2020	physics	5	essay	25	.770	Content & Construct
28	Wardhani et al.	2020	physics	5	essay	25	.800	Content & Construct
29	Faturrahman et al.	2016	chemistry	10	essay	37	.730	Content
30	Harta	2017	chemistry	5	essay	127	.930	Content & Construct
31	Hidayat et al.	2018	chemistry	4	essay	42	.710	Content
32	Ridhwan et al.	2020	biology	10	essay	136	.876	Content & Construct

Note: The reliability value shown in the table using Cronbach-alpha reliability. \*the past reliability cannot be considered as high (e.g. above .80) or at least marginally acceptable (e.g. above .60) (Gliner et al., 2017).

### ***Problem solving test framework***

In developing a problem-solving test, each study used different problem-solving frameworks and indicators. Most of the studies used Polya's problem-solving framework in which the items represent skills to understand the problem, devise a plan, carry out the plan, and evaluate the problem solution. It recorded that 11 of 32 studies used Polya's problem-solving framework and they were translated into one or more different questions. The other 3 studies used Doctor and Heller's problem-solving framework(year?) that has a specific term in physics or science context. 3 studies used OECD cognitive problem-solving framework (OECD, 2014) that was modified to fit specific mathematics and chemistry questions. Moreover, 9 studies used different references for constructing problem-solving test items. They modified concepts from different resources (Sumarmo, 2015; Butterworth & Thwaites, 2013; Jonassen, 2010; Brookhart & Nitko, 2014; Bransford et al., 1986). There is one study that did not explicitly mention their main references but only gave an explanation that used a problem-solving indicator in choosing a strategy to solve mathematics problem-solving. The last 5 studies did not explain their construct framework explicitly and only mentioned 'solving the problems'. The detailed problem-solving framework used in the literature study is shown in Table 2.

Table 2. The Problem-solving framework used in the literature study

Authors	Year	Problem-solving aspects
Alfika et al.	2018	(a) understanding the problem, (b) devising a plan, (c) carrying out the plan, and (d) evaluating the problem solution (Polya, 2004)
Anggriani et al.	2018	
Hendriyana et al.	2018	
Irawati et al.	2018	
Lestari et al.	2019	
Nadapdap et al.	2017	
Pardimin et al.	2017	
Pratiwi et al.	2015	
Subekti et al.	2014	
Wahyuningrum	2014	
Wardhani et al	2020	(a) visualizing the problem, (b) using physics approach, (c) making a specific application in physics, (d) using a mathematical procedure, and (e) defining a logic conclusion (Docktor & Heller, 2009)
Hidayat et al.	2017	
Syifaulyah Savitri et al.	2019 2020	
Bidasari	2017	(a) exploring and understanding, (b) representing and formulating, (c) planning and executing, (d) monitoring and reflecting (OECD, 2014)
Hidayat et al.	2018	
Novita	2012	
Putra	2017	(a) Formulating the problem, (b) defining semi-structural information, (c) figuring out solutions (Sumarmo, 2015)
Kurniawan et al.	2018	(a) combining imaginative skills, (b) developing modesl, (c) conducting an investigation, (d) analyzing data and arriving at a conclusion (Butterworth, J., & Thwaites, 2013)
Sinensis et al.	2019	(a) problem schema, (b) analogy, (c) causal, (d) argumentation (Jonassen, 2010)
Luthfi et al.	2019	(a) identifying the problem, (b) determining goals, (c) making strategy, (d) exploring strategy, (e) making action (Brookhart, S. M., & Nitko, 2014)
Peranginangin et al.	2019	
Sutiadi et al.	2016	
Faturrahman et al.	2016	(a) identify the problem, (b) define and represent the problem, (c) explore possible strategies, (d) act on the strategies, (e) look back and evaluate the effects of your activities (Bransford et al., 1986)
Ridhwan et al.	2020	
Yulianti et al.	2018	(a) Scientific approach, (b) structured and unstructured manner, (c) memory-based approach, and (d) no clear approach (Walsh, et al., 2007)
Februarini et al.	2017	Choosing a strategy to solve the problem
Ariawan	2016	Solving the problem
Harta	2017	
Rifa'i et al.	2018	
Sinaga	2016	
Zulkarnain et al.	2018	

### Topic distribution

The problem-solving test developed in the Indonesian context strictly follows the national curriculum regulation and administered in light of the core competencies in the curriculum. The tests addressed the specific topics and targeted different grades. From 32 studies gathered in this review, they specifically addressed a certain grade. Starting from middle school from grade 7 to grade 9, 10 studies are found for mathematics problem solving test and 4 studies for science problem solving test. In the high school, the problem-solving tests are available for mathematics (4 studies), physics (4 studies), chemistry (3 studies), and biology (1 study). Then for undergraduate level, the problem-solving tests are only found in physics (5 studies) and mathematics (1 study) courses.

The topic used in the problem-solving test is varied across the grade. The sub-topic and

subject competency are different from grades 7 to 12. For the undergraduate level, there is no restriction because every higher educational institution can modify and/or make its curriculum. That is to say, there is a general standard for higher education in Indonesia provided by the national education department, but every institution or university has the authority to make and implement its educational system. The detailed topics of problem-solving items are given in Table 3.

Table 3. Topic distribution of problem-solving assessment tools in Indonesia

Grade	Mathematics	Science		
7	Set and number, linear function, geometry	Environmental issues		
8	Statistic and probability, linear function, geometry	Force and motion, wave and sound		
9	Set and number, geometry, algebra, statistics	-		
	Mathematics	Physics	Chemistry	Biology
10	Trigonometry	Work and energy	Periodic table and chemical reaction	Environment and pollution
11	Function	Wave and sound, elasticity, optical instruments, static fluid, temperature, and heat	Stoichiometry; Acids and bases	-
12	Set and number	-	-	-
Undergraduate	Calculus	Thermodynamics; electricity; momentum and impuls	Dynamic	-

### ***Validity and reliability results of the developed tests***

This review also investigates the empirical validity and reliability tests used in the study. The developed tests used Cronbach-alpha reliability to check the item consistency and it was done by SPSS software. From the reliability result, only 3 tests have a low-reliability score ( $r < .60$ ). 19 tests have acceptable reliability score ( $r$ -value between  $.60$  and  $.80$ ) and 10 studies reported high-reliability score ( $r > .80$ ). In the term of validity, 10 studies only reported the content validity result and 22 studies reported both content and construct validity. The content validity conducted by the studies focused on the topic consistency in which the item's writing is correct based on the knowledge background and the language composition. The number of experts involved in the content validation varied from 2 to 7 people. Many of them are university and high school teachers that have experience in teaching related subjects. All studies took into consideration expert evaluations and make judgments based on these. Thus, some revisions were made until all items are considered valid by the experts.

The construct validity was done by 22 studies that showed valid evidence of the developed items. Most studies, 19 of them, used Pearson correlation analysis to measure the item's validity. They showed validity with good results, high positive  $r$  value, and significant statistics. The  $r$ -value varied between studies, the lowest value is about  $r = .380$  the highest is  $r = .880$ . All tests in the studies are considered acceptable and valid. Moreover, 3 studies used Rasch analysis in determining the validity of the item. They measure INFIT MNSQ to check the fitting items with the model. The tests showed the INFIT MNSQ index range from 0.99 to 1.03 with an acceptable range is usually from 0.7 to 1.3 (Griffin, 1999). That result means all items measure the problem-solving skills correctly.

All developed problem-solving tests in Indonesia are in the scope of certain content

knowledge and subject (mathematics, physics, chemistry, biology, and science for junior high school). The acquired knowledge from experience to the current situation will help the individual to generate a solution. Depending on the nature of a problem, different type of knowledge plays a different role in the problem-solving process because problems occur with different conditions (Liao, 2002). Because of its preference to a specific context, this assessment gives a good contribution to promoting problem-solving skills in educational practice. The tests are designed as part of teaching and learning process and the context of the test is matched with the core competency of curriculum through which the students learn. This test also serves as an evaluation of a student's competency both in problem solving skills and subject course.

Problem-solving tests are available for high school and university students. It is seen that all tests were designed for specific domains with most of them using Polya's problem-solving framework. In this case, Polya's problem solving that originally developed in 1963 is constructed for mathematics education. He claimed that the knowledge in mathematics is obtainable by using the suitable problem situation and rediscovery is a useful tool for active learning (Voskoglou, 2011). By the time, this framework has been reshaped and begun to be used not only for mathematics problem solving but also for science contexts. Moreover, other important issues related to problem solving framework are the uses of general-domain problem solving framework in specific-domain problem solving test construction. Some studies used PISA problem solving framework for mathematics and chemistry problem solving test. In PISA test of problem solving, the construct of their frameworks is implemented in the general task and not connected to any curriculum subject (i.e., traffic, climate control, and robot cleaners). The idea of using domain-general framework in specific-domain problem solving rises the universality of knowledge and principles in assessment studies.

The other studies used different problem-solving framework, but they share similar principles. For example, the first step in the process of problem-solving in every study is related to understanding the problem. It is a common task when someone tries to solve the problem. They need to know what the problem is, what variables are related to the problem, and how they understand the problem. Knowing the problem will lead to a clear-thinking path and direction for stepping in the way to the solution. Secondly, it is making the strategy of a plan for solving the strategy. All of the tests implement this aspect. Some of them put different and/or additional skills like determining goals and gathering information before making a strategy. The next steps are acting for the strategy or executing the plans they have made for the solution to the problem, and the last is making an evaluation or reflection based on the solution impact. However, the literature background used by one test developer is different from the others. Despite determining the process of making a solution, the problem-solving framework constructed by Jonassen (2010) is mentioning the analogy, causal relationship, and argumentation that are closer to being a mental process. Those skills are important as individual thinking skills in finding a solution for a given problem.

Even though the test administered in light of the topic in the curriculum, each grade is covered by the developed tests. For subjects like mathematics, it can be found the developed problem-solving test in every grade ranged from middle school to undergraduate level. Then for physics, there is only the need test for twelve grade students and in science subject there exists the test for the ninth grade. However, only few problem-solving tests were developed in chemistry (grade 10<sup>th</sup> and 11<sup>th</sup>) and biology subjects (grade 10<sup>th</sup> in environmental and pollution topic). The quite interesting thing here is about chemistry and biology problem





solving because those subjects have many connections to real “problematic” situations. Many aspects in chemistry and biology represent problem situations, for example, in chemical contamination, conservation, and pandemic disease are the real problems need to be solved in recent day. Then, introducing students these chemistry and biology-related problems through learning and assessment will them to be future problem solvers.

In the test empirical analysis, some items show acceptable Cronbach-alpha reliability. Only three developed tests have low  $r$  value. In cases where the test that has low reliability, it is better if they do some modifications or revisions to those items then re-run the analysis. Moreover, for essay type test, the scoring depends on the strength of rubric and the raters, thus it is suggested to use an additional reliability test to check the consistency of the items based on rating system such as interrater reliability (Gliner et al, 2017).

In the test validity, nearly half of the tests only conducted content validity which documented the relation of test specification with their content (Downing & Haladyna, 1997). That makes the test only checked by the expert and being validated by a personal judgment. They checked the domain used to constitute the construct. It did not show the empirical result that reflects items position based on the test taker’s perspective. The empirical investigation is critical for high stakes examination such as problem-solving skills, in order to make sure that the items correctly measure students’ skills. the other tests that used both content and construct validity, showed more solid results because they used empirical data that refers to what extent the items measure the construct (Westen & Rosenthal, 2003).

The studies used different sample numbers, but all are with relatively small numbers (21 studies used less than 50 samples, 6 studies with 50-100 samples and 5 studies with more than 100 samples). Some scholars underpin that the more samples are used in the validity measure, the better results in research quality are obtained. Schumacker & Lomax2014 also mentioned that the sample numbers for conducting validity are disparately ranging from 150 to 1000 samples based on the estimated parameter method and data normality. However, there is no exact number of samples or participants that must be included in the item's development study, in fact it depends on many factors.it will better if the sample number is relatively high and represents the exact population.

## Summary

The problem-solving skills have a beneficial impact in real life. Thus, they should be introduced early to the young generation and become a focus for educational purposes. The implementation of problem-solving skills in educational practice can be done in many aspects especially in the assessment process. In Indonesia, it was reported that 32 studies focused on problem-solving test development in specific topic based on curriculum and they are projected for high school to higher education level. The topic distribution is mostly found in mathematics (15 studies), physics (9 studies), integrated science for junior high school (4 studies), chemistry (3 studies), and biology (1 study). There are a lot of frameworks used for developing problem-solving tests in Indonesia. The most frequently used framework is Polya’s problem solving (34.4%) that was originally developed for mathematics problem solving and then reconstructed to fit into different subjects. The result reported 90% of the studies have moderate to good reliability value ( $r > .60$ ) and only 10% of them has low reliability ( $r < .60$ ). The validity analysis has been an issue since 10 studies did not report the construct validity. among 22 studies that reported construct validity, 86.4% of them performed Pearson correlation analysis and 13.6% used Rasch analysis. Apparently advance

analysis techniques for test development such as factor analysis and Rasch analysis are not widely used in Indonesian studies. Of All the studies in problem solving assessment available in the Indonesian context these are few in the term of quality and quantity. There are some limitations in the topic distribution and empirical test analysis (validity and reliability result). To this end further research concentrating on t problem-solving test development with a good research setting is needed to improve the quality of problem-solving assessment in Indonesia.

With a view to improving problem-solving research in assessment and test development applying advance empirical test analysis such as factor analysis (exploratory or confirmatory factor analysis) and Rasch analysis can be helpful. The use of a large sample number is required to get adequate statistical analysis and strong validity results. Since the test is embedded in the specific subject such as mathematics and science, more topics should be administered by the test. The review result shows that the tests dominantly available are in mathematics and physics subject. Thus, the research and the development of problem-solving assessment is urgently needed in different topics like science, chemistry, and biology targeting different educational levels.

## References

- Alfika, Z. A., Mayasari, T., & Kurniadi, E. (2019, January). Pengembangan Instrumen Tes Kemampuan Memecahkan Masalah Fisika Berbasis Kearifan Lokal pada siswa SMP/MTs [The development of physics problem solving test based on traditional culture in junior high school]. In *Prosiding SNPF (Seminar Nasional Pendidikan Fisika)*.
- Angriani, A. D., Nursalam, N., Fuadah, N., & Baharuddin, B. (2018). Pengembangan Instrumen Tes untuk Mengukur Kemampuan Pemecahan Masalah Matematika Siswa [The development of instrumental test to measure student's mathematics problem solving ability]. *AULADUNA: Jurnal Pendidikan Dasar Islam*, 5(2), 211-223. <https://doi.org/10.24252/auladuna.v5i2a9.2018>
- Ariawan, R. (2016). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Dan Komunikasi Matematis Pada Materi Bangun Ruang Sisi Datar [The development of problem solving test and mathematical communication in geometry subject]. *AdMathEdu*, 6(2), 57206. <http://dx.doi.org/10.12928/admathedu.v6i2.5447>
- Asyari, M., al Muhdhar, M. H. I., Susilo, H., & . I. (2016). Improving critical thinking skills through the integration of problem based learning and group investigation. *International Journal for Lesson and Learning Studies*, 5(1), 36-44. <https://doi.org/10.1108/IJLLS-10-2014-0042>
- Bartley, J. E., Boeving, E. R., Riedel, M. C., Bottenhorn, K. L., Salo, T., Eickhoff, S. B., Brewe, E., Sutherland, M. T., & Laird, A. R. (2018). Meta-analytic evidence for a core problem solving network across multiple representational domains. *Neuroscience & Biobehavioral Reviews*, 92, 318-337. <https://doi.org/10.1016/j.neubiorev.2018.06.009>
- Bransford, J., Sherwood, R., Vye, N., & Rieser, J. (1986). Teaching thinking and problem solving: Research foundations. *American Psychologist*, 41(10), 1078-1089. <https://doi.org/10.1037/0003-066X.41.10.1078>
- Bidasari, F. (2017). Pengembangan Soal Matematika Model PISA pada Konten Quantity untuk Mengukur Kemampuan Pemecahan Masalah Matematika Siswa Sekolah Menengah Pertama [the development of mathematic PISA test in quantity topic to measure student's mathematic problem solving skills in junior high school]. *Jurnal Gantang*, 2(1), 63-77. DOI: <https://doi.org/10.31629/jg.v2i1.59>

- Brookhart, S. M., & Nitko, A. J. (2014). *Educational assessment of students*. Pearson Higher Ed.
- Butterworth, J., & Thwaites, G. (2013). *Thinking skills: Critical thinking and problem solving*. Cambridge University Press.
- Care, E., Scoular, C., & Griffin, P. (2016). Assessment of Collaborative Problem Solving in Education Environments. *Applied Measurement in Education*, 29(4), 250–264. <https://doi.org/10.1080/08957347.2016.1209204>
- Carlson, M. P., & Bloom, I. (2005). The Cyclic Nature of Problem Solving: An Emergent Multidimensional Problem-Solving Framework. *Educational Studies in Mathematics*, 58(1), 45–75. <https://doi.org/10.1007/s10649-005-0808-x>
- Christ, A., Becker, N., & Kröner, S. (2020). Multiple complex problem-solving scenarios: The incremental validity of ability self-concept beyond reasoning in adults. *Intelligence*, 78, 101421. <https://doi.org/10.1016/j.intell.2019.101421>
- Csapó, B., & Funke, J. (2017). *The Nature of Problem Solving* (B. Csapó & J. Funke, Eds.). OECD. <https://doi.org/10.1787/9789264273955-en>
- Davidson, J. E., Sternberg, R. J., & Sternberg, R. J. (2003). *The psychology of problem solving*. Cambridge University Press.
- Dermitzaki, I., Leondari, A., & Goudas, M. (2009). Relations between young students' strategic behaviours, domain-specific self-concept, and performance in a problem-solving situation. *Learning and Instruction*, 19(2), 144–157. <https://doi.org/10.1016/j.learninstruc.2008.03.002>
- Dindar, M. (2018). An empirical study on gender, video game play, academic success and complex problem solving skills. *Computers & Education*, 125, 39–52. <https://doi.org/10.1016/j.compedu.2018.05.018>
- Docktor, J., & Heller, K. (2009). Robust assessment instrument for student problem solving. *In Proceedings of the NARST 2009 Annual Meeting, Garden Grove, CA*, 1–9.
- Downing, S. M., & Haladyna, T. M. (1997). Test Item Development: Validity Evidence From Quality Assurance Procedures. *Applied Measurement in Education*, 10(1), 61–82. [https://doi.org/10.1207/s15324818ame1001\\_4](https://doi.org/10.1207/s15324818ame1001_4)
- Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology*, 75(2), 157–169. <https://doi.org/10.1348/000709904X23411>
- Fathurrahman, M., Permanasari, A., & Siswaningsih, W. (2016). Pengembangan Tes Keterampilan Problem Solving Siswa SMA pada Pokok Bahasan Stoikiometri Larutan [The development of problem solving ability test in senior high school with stoikiometri topic]. *EduChemia (Jurnal Kimia dan Pendidikan)*, 1(1), 62-75. <http://dx.doi.org/10.30870/educhemia.v1i1.440>
- Februarini, Y. T., & Suratman, D. Pengembangan perangkat tes untuk mengukur kemampuan problem solving dan penalaran matematis siswa sekolah menengah kejuruan [The development of test series to measure student's mathematics problem solving and cognitive test in vocational high school]. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, 6(5). (retrieved from <http://jurnal.untan.ac.id/index.php/jpdpb/article/view/19934>)
- Ferreira, M., & Trudel, A. (2012). The Impact of Problem-Based Learning (PBL) on Student Attitudes Toward Science, Problem-Solving Skills, and Sense of Community in the Classroom. *The Journal of Classroom Interaction*, 47(1), 23–30.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning*, 7(1), 69–89. <https://doi.org/10.1080/13546780042000046>
- Funke, J. (2010). Complex problem solving: a case for complex cognition? *Cognitive Processing*, 11(2), 133–142. <https://doi.org/10.1007/s10339-009-0345-0>

- Glaser, R., Chi, M. T. H., & Farr, M. J. (2009). *Measuring complex problem solving: The MicroDYN approach*.
- Gliner, J.A., Morgan, G.A., & Leech, N.L. (2017). *Research method in applied settings: an integrated approach to design and analysis*. New York: Routledge, Taylor and Francis
- Gok, T. (2014). Peer instruction in the physics classroom: Effects on gender difference performance, conceptual learning, and problem solving. *Journal of Baltic Science Education*, 13(6), 776.
- Greiff, S. (2012). Assessment and Theory in Complex Problem Solving - A Continuing Contradiction? *Journal of Educational and Developmental Psychology*, 2(1). <https://doi.org/10.5539/jedp.v2n1p49>
- Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015). Assessing complex problem-solving skills with multiple complex systems. *Thinking & Reasoning*, 21(3), 356–382. <https://doi.org/10.1080/13546783.2014.989263>
- Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, 36, 37–48. <https://doi.org/10.1016/j.lindif.2014.08.003>
- Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity of complex problem solving measures. *Intelligence*, 50, 100–113. <https://doi.org/10.1016/j.intell.2015.02.007>
- Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Graesser, A. C., & Martin, R. (2014). Domain-general problem solving skills and education in the 21st century. *Educational Research Review*, 13, 74–83. <https://doi.org/10.1016/j.edurev.2014.10.002>
- Griffin, P. (1999). *Item response modeling: an introduction to Rasch model*. Melbourne: Assessment Research Center, University of Melbourne.
- Harta, J. (2017). Pengembangan Soal Esai Berbasis HOTS Untuk Menyelidiki Keterampilan Pemecahan Masalah Siswa SMA [The development of essay test based on HOTS to investigate student's problem solving ability]. *Jurnal Penelitian*, 21(1). (retrieved from <https://e-journal.usd.ac.id/index.php/JP/article/view/896>)
- Hendriana, H., Johanto, T., & Sumarmo, U. (2018). The Role of Problem-Based Learning to Improve Students' Mathematical Problem-Solving Ability and Self Confidence. *Journal on Mathematics Education*, 9(2), 291-300.
- Hidayat, S. R., Setyadin, A. H., Hermawan, K., Kaniawati, I., Suhendi, E., Siahaan, P., & Samsudin, A. (2017). Pengembangan instrumen tes keterampilan pemecahan masalah pada materi getaran, gelombang, dan bunyi [The development of problem solving skill test in the sound and wave topic]. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 3(2), 157-166. <http://doi.org/10.21009/1>
- Hidayat, T., Susilaningih, E., & Kurniawan, C. (2018). The effectiveness of enrichment test instruments design to measure students' creative thinking skills and problem-solving. *Thinking Skills and Creativity*, 29, 161-169. <https://doi.org/10.1016/j.tsc.2018.02.011>
- Hung, C. M., Hwang, G. J., & Huang, I. (2012). A project-based digital storytelling approach for improving students' learning motivation, problem-solving competence and learning achievement. *Journal of Educational Technology & Society*, 15(4), 368–379.
- Irawati, T. N., & Mahmudah, M. (2018). Pengembangan Instrumen Kemampuan Berpikir Analisis Siswa SMP dalam Menyelesaikan Soal Pemecahan Masalah Matematika [The development of student's critical thinking skills test to solve mathematics problem solving]. *Kadikma*, 9(2), 1-11.
- Iswandari, D. C., Prayogo, J. A., & Cahyono, B. Y. (2017). Effect of Environmental Problem-based Learning on the Indonesian EFL Students' Environment-related Vocabulary



- Mastery and Writing Ability. *Theory and Practice in Language Studies*, 7(8), 608. <https://doi.org/10.17507/tppls.0708.02>
- Jonassen, D. H. (2010). *Learning to solve problems: A handbook for designing problem-solving learning environments*. Routledge.
- Kroner, S., Plass, J., & Leutner, D. (2005). Intelligence assessment with computer simulations. *Intelligence*, 33(4), 347–368. <https://doi.org/10.1016/j.intell.2005.03.002>
- Kurniawan, B. R., & Taqwa, M. R. A. (2018). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Fisika pada Materi Listrik Dinamis [The development of physics problem solving test for dynamic electricity topic]. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 3(11), 1451-1457. <http://dx.doi.org/10.17977/jptpp.v3i11.11761>
- Lachman, R., Lachman, J. L., & Butterfield, E. C. (2015). *Cognitive psychology and information processing: An introduction*. Psychology Press.
- Lestari, P. E., Purwanto, A., & Sakti, I. (2019). Pengembangan instrument tes keterampilan pemecahan masalah pada konsep usaha dan energy di SMA [Th development of problem solving skills test in force and energy topic for senior high school]. *Jurnal Kumparan Fisika*, 2(3), 161-168. <https://doi.org/10.33369/jkf.2.3.161-168>
- Liao, S. H. (2002). Problem solving and knowledge inertia. *Expert Systems with Applications*, 22(1), 21–31.
- Luthfi, I. A., Muharomah, D. R., Ristanto, R. H., & Miarsyah, M. (2019). Pengembangan tes kemampuan pemecahan masalah pada isu pencemaran lingkungan [the development of problem solving test in the pollution topic]. *Jurnal BIOEDUIN: Program Studi Pendidikan Biologi*, 9(2), 11-20. <https://doi.org/10.15575/bioeduin.v9i2.5892>
- Mukhopadhyay, R. (2013). Problem solving in science learning- some important consideration of teacher. *IOSR Journal of Humanities and Social Science (IOSR-JHSS)*, 4(6), 21–25.
- Nadapdap, A. T. Y., & Istiyono, E. (2017). Developing physics problem-solving skill test for grade X students of senior high school. *REiD (Research and Evaluation in Education)*, 3(2), 114-123. <https://doi.org/10.21831/reid.v3i2.14982>
- Nickles, T. (1988). Questioning and problems in philosophy of science: Problem-solving versus directly truth-seeking epistemologies. *Questions and Questioning*, 43–67.
- Novita, R. (2012). Exploring Primary Student's Problem-Solving Ability by Doing Tasks Like PISA's Question. *Indonesian Mathematical Society Journal on Mathematics Education*, 3(2), 133-150. (retrieved from <https://eric.ed.gov/?id=EJ1078578>)
- OECD. (2014). *PISA 2012 Results: Creative Problem Solving (Volume V)*. OECD. <https://doi.org/10.1787/9789264208070-en>
- OECD. (2016). *PISA 2015 Results (Volume I)*. OECD. <https://doi.org/10.1787/9789264266490-en>
- OECD. (2017). *PISA 2015 Results (Volume V)*. OECD. <https://doi.org/10.1787/9789264285521-en>
- Pardimin, P., Widodo, S. A., & Purwaningsih, I. E. (2017). Analisis Butir Soal Tes Pemecahan Masalah Matematika [The analysis of mathematics problem solving test items]. *Wacana Akademika: Majalah Ilmiah Kependidikan*, 1(1). <http://dx.doi.org/10.30738/wa.v1i1.1084>
- Partnership for 21<sup>st</sup> century skills (2009). *P21 Framework Definitions*. (retrieved from <https://files.eric.ed.gov/fulltext/ED519462.pdf>)
- Peranganing, S. A., & Surya, E. (2017). An analysis of students' mathematics problem solving ability in VII grade at smp negeri 4 pancurbatu. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 33(2), 57-67.
- Polya, G. (1945). *How to solve it*. Princeton University Press.



- Polya, G. (2004). *How to solve it: A new aspect of mathematical method*. Princeton University Press.
- Pratiwi, D.N.. (2015). Pengembangan Instrumen Evaluasi Berbasis Taksonomi Structure of the Observed Learning Outcome (SOLO) Untuk Menentukan Profil Kemampuan Siswa dalam Memecahkan Masalah Fluida Statis [The development of evaluation instruments based on the Structure of the Observed Learning Outcome (SOLO) taxonomy to measure students problem solving skills in static fluid topic]. *Inovasi Pendidikan Fisika*, 4(3).
- Putra, H. D. (2017). Pengembangan Instrumen untuk Meningkatkan Kemampuan Mathematical Problem Posing Siswa SMA [The development of instrument to foster student's mathematical problem posing in senior high school]. *Euclid*, 4(1). DOI: <http://dx.doi.org/10.33603/e.v4i1.211>
- Rausch, A., & Wuttke, E. (2016). Development of multi faceted model of domain specific problem solving and its acceptance by different stakeholders in the business domain. *Unterrichtswissenschaft*, 44(2), 169–184.
- Ridhwan, R., Sumarmi, S., Ruja, I., Utomo, D., & Sari, R. (2020). Measuring Students Environmental Problem Solving Ability Across Gender and School Differences Using Paper Based Testing. *International Journal of Emerging Technologies in Learning (IJET)*, 15(13), 303-320.
- Rifa'i, A., & Lestari, H. P. (2018, March). The effect of think pair share (TPS) using scientific approach on students' self-confidence and mathematical problem-solving. In *Journal of Physics: Conference Series*. <https://doi.org/10.1088/1742-6596/983/1/012084>.
- Rudolph, J., Niepel, C., Greiff, S., Goldhammer, F., & Kröner, S. (2017). Metacognitive confidence judgments and their link to complex problem solving. *Intelligence*, 63, 1–8. <https://doi.org/10.1016/j.intell.2017.04.005>
- Savitri, N., & Setiani, I. A. (2020). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Fisika pada Materi Momentum dan Impuls [The development of problem solving test in physics: momentum and impuls]. *Jurnal Kependidikan Betara*, 1(1), 20-26.
- Schoenfeld, A.H. (2013). Reflections on problem solving theory and practice. *The Mathematics Enthusiast*, 10(1), 9–35. <https://scholarworks.umt.edu/tme/vol10/iss1/3>
- Schoenfeld, Alan H. (1987). Pólya, Problem Solving, and Education. *Mathematics Magazine*, 60(5), 283–291. <https://doi.org/10.1080/0025570X.1987.11977325>
- Schumacker, R. E., & Lomax, R. G. (2014). *A beginner's guide to structural equation modeling*. Psychology Press.
- Sinaga, N. A. (2016). Pengembangan tes kemampuan pemecahan masalah dan penalaran matematika siswa SMP kelas VIII [The development of student's mathematics problem solving and cognitive test in 8<sup>th</sup> grade]. *Pythagoras: Jurnal Pendidikan Matematika*, 11(2), 169-181. <https://doi.org/10.21831/pg.v11i2.10642>
- Sinensis, A. R., Firman, H., Hamidah, I., & Muslim, M. (2019). Pengembangan instrument tes termodinamikan untuk mengukur kemampuan pemecahan masalah (PsACAr) pada mahasiswa calon Guru Fisika [The development of test in thermodymanic to assess preservice teacher's ability on problem solving]. *Jurnal Inovasi dan Pembelajaran Fisika*, 6(2), 122-129. <https://doi.org/10.36706/jipf.v6i2.9879>
- Sonnleitner, P., Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R., Hazotte, C., Mayer, H., & Latour, T. (2012). The Genetics Lab. Acceptance and psychometric characteristics of a computer-based microworld to assess complex problem solving. *Psychological Test and Assessment Modeling*, 54, 54–72.

- Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, 53, 92–101. <https://doi.org/10.1016/j.intell.2015.09.005>
- Sternberg, R. (2018). Speculations on the Role of Successful Intelligence in Solving Contemporary World Problems. *Journal of Intelligence*, 6(1), 4. <https://doi.org/10.3390/jintelligence6010004>
- Subekti, F. E., & Widiyanti, R. A. (2014). Pengembangan instrumen tes Kalkulus Lanjut 2 berbasis pemecahan masalah [The development of advance calculus test based on problem solving framework]. *Khazanah Pendidikan*, 7(1). DOI: 10.30595/jkp.v7i1.669
- Sugrue, B. (2005). A Theory-Based Framework for Assessing Domainl-Specific Problem-Solving Ability. *Educational Measurement: Issues and Practice*, 14(3), 29–35. <https://doi.org/10.1111/j.1745-3992.1995.tb00865.x>
- Sumarmo, U. (2015). *Mathematical Problem Posing: Rasional, Pengertian, Pembelajaran, dan Pengukurannya* [Mathematical problem solving: rational, definition, learning, and assessment]. Pascasarjana STKIP Siliwangi Bandung dan Pascasarjana UPI.
- Sutiadi, A., & Nurwijayaningsih, H. (2016). Konstruksi dan Profil Problem Solving Skill Siswa SMP dalam Materi Pesawat Sederhana [The construction and profile of student's problem solving skills for simple machine topic in junior high school]. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, 2(1), 37-42. <https://doi.org/10.21009/1.02106>
- Syifaulyah, A. (2019, November). Pengembangan Instrumen Tes Kemampuan Memecahkan Masalah Fisika Bunyi Berbasis Musik [The development of problem solving in music based sounds topic]. In *Seminar Nasional Lontar Physics Forum* (pp. 100-105).
- Voskoglou, M. G. (2011). Problem solving from Polya to nowadays: A review and future perspectives. *Progress in Education*, 22(4), 65–82. [https://d1wqtxts1xzle7.cloudfront.net/32888949/Problem\\_-\\_Solving.pdf?1391217514=&response-content-disposition=inline%3B+filename%3DProblem\\_Solving\\_from\\_Polya\\_to\\_Nowadays\\_A.pdf&Expires=1598531286&Signature=GK3T3jRs2cyK-Cw8IYirLoBEeC51q10f1eTT3qSDhzEmbHouDK](https://d1wqtxts1xzle7.cloudfront.net/32888949/Problem_-_Solving.pdf?1391217514=&response-content-disposition=inline%3B+filename%3DProblem_Solving_from_Polya_to_Nowadays_A.pdf&Expires=1598531286&Signature=GK3T3jRs2cyK-Cw8IYirLoBEeC51q10f1eTT3qSDhzEmbHouDK)
- Wahyuningrum, E. (2014). *Kemampuan pemecahan masalah, komunikasi dan disposisi matematik siswa SMP melalui strategi MEAs* [The mathematics problem solving skills, communication and disposition skills in junior high schools based on MEAs strategy]. Dissertation. Unpublish. Graduate program Universitas Pendidikan Indonesia
- Walker, F., Link, F., & Nickolaus, N. (2016). A multidimensional structure of domain-specific problem-solving competencies of electronics technicians for automaton technology. *Empirical Research in Vocational Education and Training*, 8(7), 1–26. <https://doi.org/10.1186/s40461-016-0034-z>.
- Walsh, L. N., Howard, R. G., & Bowe, B. (2007). Phenomenographic study of students' problem solving approaches in physics. *Physical Review Special Topics-Physics Education Research*, 3(2), 020108. <https://doi.org/10.1103/PhysRevSTPER.3.020108>
- Wardani, A., Arkan, M. N., & Suyudi, A. (2020). Pengembangan Instrumen Tes Keterampilan Pemecahan Masalah Pokok Bahasan Listrik Dinamis [The development of problem solving test in dynamic electricity]. *Jurnal Kependidikan Betara*, 1(1), 14-19.
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, 84(3), 608–618. <https://doi.org/10.1037/0022-3514.84.3.608>

- Wolff, K. (2017). Engineering problem-solving knowledge: the impact of context. *Journal of Education and Work*, 30(8), 840–853. <https://doi.org/10.1080/13639080.2017.1380299>
- Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving — More than reasoning? *Intelligence*, 40(1), 1–14. <https://doi.org/10.1016/j.intell.2011.11.003>
- Yu, W., She, H., & Lee, Y. (2010). The effects of Web-based/non-Web-based problem-solving instruction and high/low achievement on students' problem-solving ability and biology achievement. *Innovations in Education and Teaching International*, 47(2), 187–199. <https://doi.org/10.1080/14703291003718927>
- Yuliati, L., Riantoni, C., & Mufti, N. (2018). Problem Solving Skills on Direct Current Electricity through Inquiry-Based Learning with PhET Simulations. *International Journal of Instruction*, 11(4), 123-138.
- Zulkarnain, R., Sugiarno, S., & Bistari, B. (2018). Pengembangan instrument problem solving dengan sajian masalah matematika kontekstual di sekolah menengah pertama [The development of problem solving instruments with mathematics contextual problems in senior high school]. *Jurnal Pendidikan dan Pembelajaran*, 7(7).