International Journal of Instruction *January 2023 ● Vol.16, No.1*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 723-740

Article submission code: 20211217051334
Received: 17/12/2021 Revision: 11/08/2022
Accepted: 05/09/2022 OnlineFirst: 11/11/2022

Innovation of Analytical Thinking Skills Instrument for Throwing and Catching Game Activities for Elementary School Students

Jusuf Blegur
Universitas Kristen Artha Wacana, Indonesia, *jusufblegur@ukaw.ac.id*

Christin Prima Mery Rajagukguk
Universitas Kristen Artha Wacana, Indonesia, *christin@ukaw.ac.id*

Alya Elita Sjioen
Universitas Kristen Artha Wacana, Indonesia, *alya.es@ukaw.ac.id*

Mieke Souisa
Universitas Pattimura, Indonesia, *ms.souisa1512@gmail.com*

Analytical thinking skills are one aspect of the high order thinking skills (HOTS) group, so teachers must internalize and engineer student learning activities to foster and improve students' HOTS through various throw–and–catch games. At the same time, physical education teachers are still limited in developing their measurement instruments. This study aims to develop the instrument for students' analytical thinking skills in throwing and catching game activities. This study used a research and development approach with 72 (M = 11.60, SD = 0.744) elementary school students as the subject to trial. Nine experts were involved in this research, six from the theoretical element and three from the practitioner/construct element. The instrument was developed using three main indicators, namely differentiating (3 questions), organizing indicators (3 questions), and attributing indicators (3 questions). The analysis used was descriptive and Aiken–V to test content validity and factor analysis and Pearson correlation to test the construct validity. Meanwhile, to test the reliability using the formula from Cronbach's alpha, the analysis of the difficulty of the questions and the differentiating power of the questions used the formula developed by Sharma. The findings of content validity, construct validity, reliability, difficulty, and discriminating power of the questions as a whole indicate that nine questions can be used to measure analytical thinking skills in the throwing and catching game for elementary school students.

Keywords: analytical thinking, higher–order thinking skills, throw–and–catch games, physical education learning, elementary school students

Citation: Blegur, J., Rajagukguk, C. P. M., Sjioen, A. E. & Souisa, M. (2023). Innovation of analytical thinking skills instrument for throwing and catching game activities for elementary school students. *International Journal of Instruction*, *16*(1), 723-740. https://doi.org/10.29333/iji.2023.16140a

INTRODUCTION

Physical education learning uses physical media to achieve educational goals. This condition often causes teachers to neglect students' thinking skills, especially higher–order thinking skills. In fact, through physical education learning activities, teachers are also responsible for improving students' higher–order thinking skills. If we view some of the development of quality instruments for learning physical education, researchers have emphasized the development of students' thinking skills. For example, in the article, Ho, Ahmed, & Kukurova (2021), have added the statement item "Helping students to develop critical thinking skills" in the indicator of developing thinking skills. Also, in Maksum' (2012) teaching performance instrument, precisely in item 12, he included the statement, "Teachers ask questions to stimulate students' thinking." Higher–order thinking skills (including analytical thinking) are competencies students need to compete and achieve success in the professional field (Ramdiah et al., 2018). Due to the various challenges in the workplace ahead and the rapid progress of innovation, higher–order thinking skills can help students diagnose problems or concepts in detail, rationalize the reasons with the indicators, and critically find the connection among indicators to help students solve problems comprehensively.

Analytical thinking skills need to be taught to students (Yulina et al., 2019), because analytical thinking skills help someone in elaborating a concept into more detailed sections and being able to explain the relationship between indicators (Laksono et al., 2017; Hasyim, 2018) or able to identify the real intention and the conclusion relationship between concepts, descriptions, or other forms of representation to express beliefs, reasons, information (Prawita, Prayitno, & Sugiyarto, 2019). Darmawan's (2020) findings prove that students' analytical thinking skills can be identified through algorithm clarity, chronological reasoning, valid arguments, and practical steps. As a result, analytical thinking skills will help students account for the various decisions they make when investigating a problem because they critically distinguish problem indicators, comprehensively organize problems, and are rational in relating to problem indicators. As a result, according to Mayarni & Nopiyanti (2021), analytical is an essential indicator in developing critical thinking skills.

Analytical thinking skills can help students solve problems and help students make wise decisions by thinking logically (Mayarni & Nopiyanti, 2021), so it is not easy to believe in superstition (Tosyali & Aktas, 2021) because it prioritizes utilitarian moral judgments (Li et al., 2018). Teachers can design effective learning activities to improve analytical thinking skills (Khairunnisa, Sutjihati, & Retnowati, 2021). For example, using learning with the vision of Science, Environment, Technology, and Society (Maghfiroh & Sugianto, 2011), developing a Visual Learning Model (Raiyn, 2016), applying Context–Based Learning (Sudibyo, Jatmiko, & Widodo, 2016), using the Chemistry Systemic Learning Approach (Fitriyana, Marfuatun, & Priyambodo, 2019), applying the Research–Oriented model Collaborative Inquiry Learning (Huda & Rohaeti, 2020), using the Electronic Collaborative model (E–CoPAL) (Sumaryati et al., 2020), developing Student Worksheets (Bierera & Muchlis, 2021), as well as optimizing the Scientific Approach equipped with Student Worksheets (Rahayu, Handoyo, & Purwito,

2021). These learning models succeeded in restoring students' analytical thinking skills. They help students with operational guidelines to facilitate exploratory and identifying thinking processes on various contextual problems so that the rationalization process behind these problems can run optimally.

Although students learn to use physical media, teachers are also responsible for measuring analytical thinking skills and students' character as a unit in determining student learning outcomes. Throw–and–catch games highlighted the psychomotor activity of students. However, students must be able to rationalize logical arguments in each of their decisions (for example, deciding to throw the ball up when their partner is far off). The problem of repairing and improving analytical thinking skills in learning has not become a serious concern for physical education teachers. It has implications for the delay in implementing various learning activities that support students to optimize student HOTS. The Anggraeni & Sole (2020) study report is astounding. They found that teachers did not understand the definition or the indicators of HOTS, so teachers were unable to develop instruments and verbs to measure students' HOTS. Instead, they measure students' lower–order thinking skills, such as remembering and explaining (Abosalem, 2016). Furthermore, Thaneerananon, Triampo, & Nokkaew (2016), Prawita, Prayitno, & Sugiyarto (2019) and Fitriani, Fadly, & Faizah (2021) also proved their studies that students' HOTS was low. Four of the six lowest indicators of student HOTS from Fitriani, Fadly, & Faizah's (2021) study are understanding concepts (3.8), identifying (3.5), organizing (3.75), and connecting (4.6). These results strengthen the previous study of Ikmah, Nugroho, & Sutikno (2018), namely the weakness of students in developing analytical thinking skills because they cannot understand the relationship between variables with each other in the case of specific problems.

Students' thinking ability shows gradations (low and high), but teachers need credible norms and indicators to justify this thinking ability. The concept table from Anderson et al. (2001) places the analysis in the higher–level thinking level of one's knowledge domain. The results of the study by Montaku (2011) confirmed that people who have high analytical thinking skills have better academic performance and learning attitudes. Even so, the study of Yuwono, Sunarno, & Aminah (2020) revealed interesting contradictory data. They found that students' analytical thinking skills had no significant effect on learning outcomes in the realm of knowledge. The difference can be motivated by instruments to measure students' analytical thinking skills, in which Yowono and colleagues used 12 multiple–choice tests to measure students' analytical thinking skills. Whereas through the essay, the analysis process is more "maximum" because students must be able to distinguish, organize, and connect a particular problem/topic using rational argumentation (Ramirez & Ganaden, 2008; Laksono et al., 2017). Through social essays, students are not only offered a "impromtu decision," but each student's answer must be able to rationalize analytically and critically to confirm the gradation of thinking processes that have the potential to support the development of science.

Despite the two previous contradictory study reports, we agree that teachers need a valid and reliable instrument to measure analytical thinking skills. If the teacher ignores the validity and reliability of the instrument, the credibility of his decision is questionable.

That is why numerous researchers tried to develop analytical thinking skills instruments according to the characteristics of their respective fields of science and learning materials. For example, Ad'hiya & Laksono (2018) developed the Integrated Assessment–Analytical Thinking and Chemical Literacy (IA–ATCL) with Aiken's validity value of 0.86. Laksono et al. (2017) developed the Analytical Thinking and Science Process Skill (ATSPS) instrument with a validity value of 0.95. Wiyarsi et al. (2019) developed a Multiple Representation–Based Analytical Thinking Test (TAT–MR). The three instruments above are applied in chemistry learning. Irawati & Mahmudah (2018) have also developed an analytical thinking skill instrument for mathematical problem–solving problems with a reliability value of 0.78. Meanwhile, Fitriani, Suryana, & Hamdu (2018) succeeded in developing the Higher–Order Thinking Skill Test in outdoor Learning–Based Thematic Learning with a Cronbach alpha reliability value of 0.935. These instruments have gone through a rigorous process of substance and methodological verification to suppress any bias caused by the subjectivity of the assessors when measuring and assessing students' thinking abilities.

Some of the snippets of developing analytical thinking skills instruments above show that there is no single analytical thinking skill instrument that specifically reviews students' throwing–catch game activities. In fact, in physical education learning, students' basic movement skills (manipulative) that they often use are through throwing and catching games. Thus, the results of this study contribute to the teacher's decision in measuring and assessing students' analytical thinking skills through throw–and–catch game activities through valid and reliable instruments. In short, even though students use their physical mobility to achieve learning objectives, teachers are still careful in diagnosing students' analytical thinking skills to support the development of holistic student learning outcomes.

METHOD

Development Model and Procedure

The model used in this instrument development research is the model of Oriondo & Antonio (1998). The instrument's stages were through a test with a modified model from Oriondo and Antonio as follows, 1) design, 2) trial, 3) test measurement.

Table 1
Analytical thinking skills instrument validator

No	Name	Descrition	Affiliation
1	Prof. Dr. I Made Sriundy Mahardika, M.Pd.	The validators of theory	Universitas Negeri Surabaya
2	Prof. Dr. Jacob Anaktototy, M.Pd.		Universitas Pattimura
3	Dr. Lukas M. Boleng, M.Kes.		Universitas Nusa Cendana
4	Dr. Wahyu Indra Bayu, M.Pd.		Universitas Sriwijaya
5	Dr. Rahayu Prasetiyo, M.Pd.		STKIP PGRI Jombang
6	Dr. Yusuf Hidayat, M.Si.		Universitas Pendidikan Indonesia
7	Yohanis Boling, S.Pd., M.Pd.	The validators of construction/practice	SD Negeri Balfai
8	Ismael Fernandez, S.Pd.		SD Inpres Naikoten I
9	Dintje Terfina Eky, S.Pd.		SD Negeri Balfai

Design and Test Subjects

This study attempted to develop a test instrument using essay questions to measure students' analytical thinking skills in throwing and catching games. There are nine essay questions, consisting of 3 questions from differentiating indicators, three questions from organizing indicators, and three questions from atributing indicators. The questions were then validated by nine experts, namely six theorists and three constructors/practical experts, descriptively to find the mean value (≥ 3) and test content validity with Aiken–V and construct validation using factor analysis and correlation analysis from Pearson. After all the questions passed the validation stage, the questions were tested on grade VI elementary school students. Each student's answers were assessed based on an assessment rubric with four scales (1–4), as shown in Table 2.

Table 2
Rubric for assessing analytical thinking skills for throwing and catching essay questions

Indicators	Grading scale			
	4	3	2	1
Differentiating (questions 1–3)	Students are able to distinguish relevant or important parts from irrelevant or unimportant parts of the material presented	Students are quite able to distinguish relevant or important parts from irrelevant or unimportant parts of the material presented.	Students are less able to distinguish relevant or important parts from irrelevant or unimportant parts of the material presented.	Students are unable to distinguish relevant or important parts from irrelevant or unimportant parts of the material presented.
Organizing (questions 4–6)	Students are able to determine how elements fit or function within a structure.	Students are quite able to determine how elements fit or function in a structure.	Students are less able to determine how elements fit or function in a structure.	Students are unable to determine how elements fit or function in a structure
Atributing (questions 7–9)	Students are able to determine the point of view, bias, value, or intent that underlies the material presented.	Students are quite able to determine the point of view, bias, value, or intent that underlies the material presented.	Students are less able to determine the point of view, bias, value, or intent that underlies the material presented.	Students are not able to determine the point of view, bias, value, or intent that underlies the material presented.

The results obtained were then analyzed for validity and reliability and looked for the level of accuracy of the questions and the differentiating power of the questions to meet the development and dissemination requirements.

The experimental subjects in this study were students of class VI Balfai State Elementary School, Kupang Regency, East Nusa Tenggara Province, Indonesia,

totalling 72 students, (male = 30 or 41.7%, female = 42 or 58.3%; M = 11.60, SD = 0.744). The determination of the test subjects was taken using a total sampling technique that is students of class VI/a–d.

Data Collection

The analytical thinking skills test used an essay test developed from 3 indicators from Anderson et al. (2001), Ramirez & Ganada (2008), and Laksono et al. (2017). The three indicators of analytical thinking skills are distinguishing skills, organizing skills, and connecting skills. In connection with the needs of the research population for elementary school students and the activity of throwing and catching games, the formulation of the questions also focuses on these two needs. These needs include the limitations of questions (9 numbers) and the scope of questions given to students. For example, test questions about when students pass the ball, the position to catch the ball, or the area to throw or catch.

Table 3
Indicators, definitions, and analytical thinking skills test essay questions

Indicators	Definition	Questions
Differentiating	Distinguish the relevant or essential part from the irrelevant or unimportant part of the material presented	1. Why do you need to catch the ball using both hands when the ball thrown by a friend is very fast? 2. How do you catch the hard ballso it does not fall? 3. How do you throw the ball to make it easy for friends to catch?
Organizing	Determine how elements fit or function within a structure	4. Why do you have to throw a high–speed ball at a friend? 5. Why must the ball be passed to another friend in the game? 6. Why do you need to look at the target or friends when you throw or catch the ball?
Atributing	Determine the point of view, bias, value, or intent that underlies the material presented	7. Why do you need the correct throwing and catching technique in the game? 8. Why do you need another member in a team? 9. Why don't you pass the ball to a friend out of reach?

Data Analysis

Descriptive analysis was used to calculate the value of the validation from the theorist, and constructor/practitioner (content validity) on the students' analytical thinking skills instrument, with the average value, received 3 of 4 interval scale. According to the validator's notes, if the expert validation value does not meet 3, the instrument will be improved both from the theoretical aspect and the construction/practical aspect.

The content validity was also tested using the Aiken–V formula (Aiken, 1985) and Inter–Rater Reliability using Pearson's Intraclass Correlation Coefficients. Meanwhile,

construct validity was tested using factor analysis and Pearson's correlation analysis. Furthermore, the instrument's reliability was tested with Cronbach alpha using the norm from Ceniza & Cereno (2012). A construct is declared reliable if it gives a value of 0.60 (Sunyoto, 2011).

To test the difficulty level of essay questions, the research team used the formula from Sharma (2021) and the norm from Thorndike & Hagen (1977), namely: 1) easy (0.70–1.0), 2) sufficient (0.30–0.70), and 3) difficult (0.00–0.30). Meanwhile, to test the discriminatory power of the questions, using the formula from Sharma (2021) and the norms of Sudijono (2011), namely: 1) poor (0.00–0.20), 2) fair (0.21–0.40), 3) good (0.40–0.70), and 4) very good (0.70–1.00). All tests used Microsoft Excel and the SPSS version 25 application.

FINDINGS AND DISCUSSION

Content Validity

Description analysis determined the distribution of the average values that can be accepted from the validator's results. In this study, the average value received was 3 of the 4 point scale. The results showed that although the validators gave various values; the highest score of 4 and the lowest value of 1. Overall, the validators agreed that this instrument could be used or had fulfilled the description analysis because all the essay questions on the analytical thinking ability of elementary school students were in the game of throwing and catching.

Table 4
Descriptive analysis of research instruments

Q	Theory validators						Constructive validator			M	SD
	1	2	3	4	5	6	7	8	9		
1	3	4	2	3	4	4	4	3	4	3.44	0.72
2	4	3	3	3	4	4	3	4	3	3.44	0.52
3	3	4	3	3	4	4	3	4	4	3.55	0.52
4	4	4	4	3	4	4	3	2	3	3.44	0.72
5	4	3	4	3	4	4	4	4	3	3.66	0.50
6	3	4	3	3	4	4	2	3	3	3.22	0.66
7	4	4	3	3	4	4	3	2	4	3.44	0.72
8	4	3	1	3	3	4	4	4	3	3.22	0.97
9	3	4	4	3	4	4	3	4	4	3.66	0.52

Instruments distributed to validators have included notes in the last column so that the validator can provide input or notes used by researchers to make improvements to enter the content validity analysis (Aiken–V and Inter–Rater Reliability) and construct stages (factor analysis and Pearson analysis).

Nine questions with a 4–point scale that the expert assessment panel has validated indicate the suitability of the test instrument to be used in the context of elementary school students' analytical thinking skills for throw–and–catch game material. The validator's response revealed that the essay question was a suitable instrument for both theories and construct. The Aiken–V results are described in Figure 1, where all items

pass the limit value of 0.74 (≥ 0.74). The value of 0.74 is obtained from the threshold value for nine validators with 4 rating categories based on the Aiken table (1985). Meanwhile, the Inter–Rater Reliability test results using Intraclass Correlation Coefficients prove that the average agreement value for validators is 0.197. As for the consistency for one person as big as 0.688..

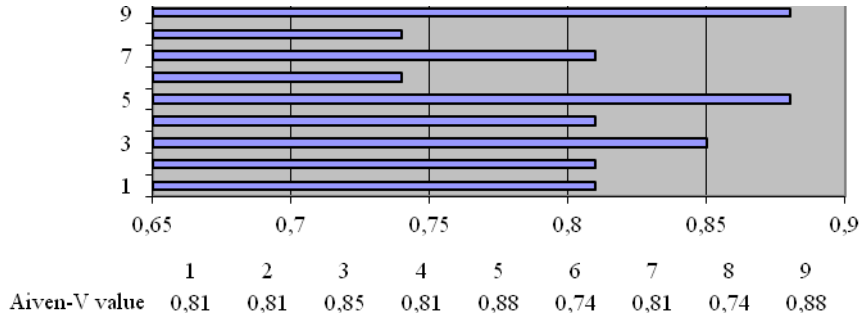

Figure 1
Aiken–V value

Even though the validation results were met (≥ 0.74), the validator provided notes that needed to be followed up before constructing validity. For example, for question number 1, the language needs to be simplified. Another complete note the different validator suggested was to recommend that the question be changed to "In what situation did you catch the ball with both hands?". Alternatively, "If the ball is thrown hard by a teammate, what will you do? Tell me your reasons!" That is, in this indicator (distinguishing), different situations must be displayed for students to be able to choose and determine the movement to be carried out by stating the reasons.

As for question number 2, the notes from the validator, including questions, also explore the reasons for the respondents' answers so that the reasons given will describe the ability to analyze respondents. Question number 4, the construct validator gives a score of 2, noting that the basic technique of throwing and catching games must be adjusted to the distance between A and B. If the distance is far, the thrower must throw a strong throw at the receiver. In addition, from the theory validators commenting that the specified indicators are related to organization, the question should begin with the word "How." That way, students will try to assemble sentences by imagining the situation of the throw–and–catch game.

For question number 7, the constructor/practitioner validator scored two because he thought the question did not fit the existing definition. While another theory validator notes that if the answer is because it is a team game, this question does not work. Continuing to question number 8, an important note from the theoretical validator is what does the sentence mean to know the game's understanding, considering that a team game requires other members? Or the reason for the team function in team play? There

is the ambiguity of understanding. It is better to clarify the question according to the desired purpose and represent the variables of analytical thinking.

In general, another addition from the validator for organizing indicators is that the specified indicator is related to linking, so the question should begin with the word "why." That way, students will try to answer by compiling situations in the throwing–and–catching game. In addition, to maintain a hierarchical teacher–student relationship, perhaps the word "you" can be replaced with "you." The words "friends," "friends," and "other friends" may be replaced by "teammates."

Based on the results of content validation, improvements were made to some questions developed before entering the trial. The improvement of the essay question instrument to measure analytical thinking skills is as follows.

Table 5
Revision of instruments (essay problems) analytical thinking skills of elementary school students in throwing and catching games

No	Before	Revision
1	Why do you need to catch the ball with both hands when your teammates throw the ball hard?	In what situation do you catch the ball with both hands?
2	How do you catch a hard ball, so it does not fall?	What will you do if the ball thrown by a teammate is fast?
3	How do you throw the ball, so it is easy for friends to catch?	How do you throw the ball so that your teammates easily catch it?
4	Why do you have to throw a hard ball at a friend?	Why do you have to throw a hard ball at a teammate?
5	Why does the ball have to be passed to another friend in the game?	Why does the ball have to be passed to another teammate in team play?
6	How do you catch a ball that is not on target?	How do you try to catch a ball thrown by a teammate when it is not on target?
7	Why do you need the correct throwing and catching technique in the game?	Why do you need the correct throwing and catching technique in the game?
8	Why do you need another member in a team game?	Why are teammates important in team sports to win the game?
9	How do you pass the ball to a friend out of reach?	How do you pass the ball to a teammate who is out of reach?

Construct Validity

Construct validity was carried out using factor analysis and Pearson analysis. Factor analysis helps reduce the number of reported variables by determining the significant variables and combining them into a single variable or factor. It can be used to find factors or possible test hypotheses (Polit & Beck, 2008).

The results of factor testing found the Kaiser–Meyer–Olkin (KMO) value of 0.709. Hair et al. (2006) viewed that the KMO value of 0.5–0.7 was moderate. While the value of 0.7–0.8 was high. Thus, the KMO score for the essay instrument was classified as good (0.709). In contrast, Bartlett's Test of Sphericity test results found a value of 119.299

with a significance value of 0.000. It means that the significance value is lower than 0.05. Thus the instrument is eligible to pass factor analysis.

After receiving the KMO and Bartlett's Test of Sphericity values, it is continued by looking at the Anti–image Correlation value. Anti–image Correlation values for the 9 question items ranged from 0.667 to 0.763. This value indicates that no factor has a value below 0.50, so the value of the question load highly contributes to the instrument's factor structure. In conclusion, this instrument has met the requirements to test unidimensional assumptions. The results of the unidimensionality assumption test are shown in Table 6.

Table 6
Total varian explained

Component	Initial eigenvalues		
	Total	% of Variance	Cumulative %
1	2.823	31.372	31.372
2	1.554	17.268	48.640
3	0.997	11.075	59.715
4	0.852	9.465	69.180
5	0.797	8.858	78.038
6	0.613	6.813	84.851
7	0.499	5.543	90.394
8	0.456	5.062	95.456
9	0.409	4.544	100.000

In the table *Total Variance Explained*, shows two factors formed from the nine indicators entered (> 1.000). Factor 1 eigenvalue of 2.823 with variance (31.372%) and factor 2 eigenvalue of 1.554 with variance (17.268%). The amount of variance that the newly formed factor can explain is 48.640%, while other factors explain the remaining 51.368%.

Figure 2
Scree plot of factor analysis

Figure 2 shows that the curve starts to flatten relative to the third factor, so it can be stated that at least two factors are formed, with the first factor being the dominant factor (48.640%). According to Reckase (1979), Smits et al. (2011), Wu et al. (2013), and Sadhu & Laksono (2018), if the factor analysis output in the factor analysis produced by the first factor can explain the variance 20%, then the unidimensionality assumption can be accepted.

Another construct validity analysis that is also used is Pearson. Pearson's test is one way to test the construct validity by comparing the total correlation value with the r–table value. A construct is said to be valid if it exceeds the r–table value. The results of the Pearson validity test are as follows.

Tabel 7
Analisis pearson

Q	Pearson correlation	r_{table}	Sig.	Decision
1	0.457	0.235	0.000	Valid
2	0.562	0.235	0.000	Valid
3	0.633	0.235	0.000	Valid
4	0.533	0.235	0.000	Valid
5	0.495	0.235	0.000	Valid
6	0.472	0.235	0.000	Valid
7	0.543	0.235	0.000	Valid
8	0.654	0.235	0.000	Valid
9	0.635	0.235	0.000	Valid

Based on Table 7, the results of the validity analysis using the Pearson correlation found that the nine items developed met the validity standard. The distribution of r values from 0.457 to 0.654 with a significance of 0.000. It means that all question items have a rcount value greater than rtable (0.235). Thus, based on the results of construct validation using Pearson's analysis, nine question items were declared valid.

Reliability

Cronbach's alpha internal consistency coefficient was found to be 0.712. It proves a 71.2% certainty of the consistency of the test items in producing more or less the same results repeatedly. As a result, essay questions to measure the analytical thinking skills of elementary school students in a throw–and–catch game are reliable to use. More specifically, if we use the norm from Ceniza & Cereno (2012), the reliability coefficient in this instrument is high because it is in the range of values from 0.61 to 0.80.

Problem Difficulty

The difficulty of the questions intends to diagnose difficult, medium, and easy questions in a construct. According to Sadhu & Laksono (2018), the difficulty index of each item is carried out to get an idea of the proportion of test–takers who answer the questions correctly. Testing the difficulty level of essay questions in this research case uses the Sharma (2021) formula and the norms from Thorndike & Hagen (1977), namely: 1) easy (0.70–1.0), 2) sufficient (0.30–0, 70), and 3) complex (0.00–0.30). The analysis results prove that 4 (44.44%) questions are classified as easy, and 5 (55.55%) questions are considered sufficient. Thus, there was no single question included in the difficult category from the nine questions constructed.

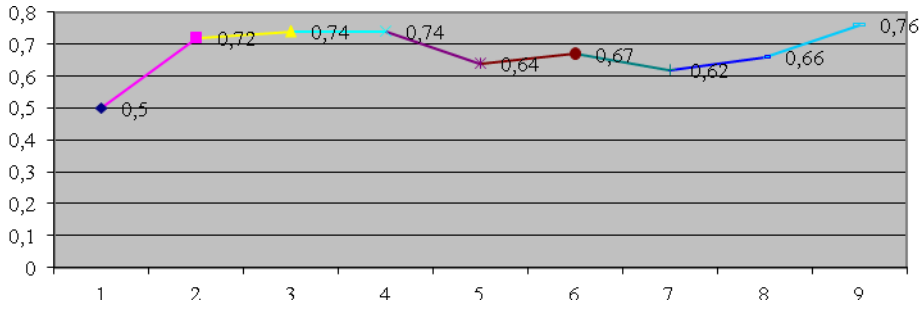

Figure 3
Question difficulty level graph

Power of Differing Questions

The discriminatory power analysis was carried out to determine whether or not an item could distinguish between high–ability participants and low–ability participants. According to Karim, Sudiro, & Sakinah (2021), by analyzing the discriminatory power of the questions, we can find out that the items effectively evaluate students' thinking

abilities. To test the discriminatory power of the questions, the formulas from Sharma (2021) and the norms of Sudijono (2011), namely: 1) poor (0.00–0.20), 2) sufficient (0.21–0.40), 3) good (0.40–0.70), and 4) very good (0.70–1.00). The test results prove that the difference between students' analytical thinking skills is only in two categories. Each category has five questions (55.55%), and the 'good' category has four questions (44.44%).

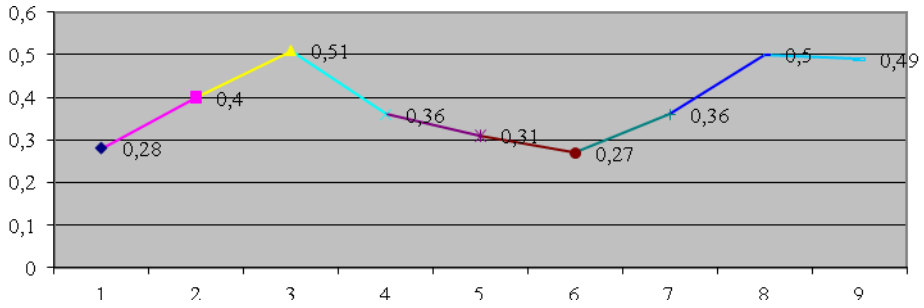

Figure 4
Graph of differential power of questions

CONCLUSION

The innovation of this instrument is not used to measure and assess the students' performance in throwing and catching the ball but rather to measure and assess students' analytical thinking skills for every cognitive decision chosen when throwing or catching the ball because every physical activity that students choose (including throwing and catching a ball) is stimulated by cognitive considerations. In addition, the assessment of learning outcomes uses cognitive, affective, and psychomotor taxonomies. Although physical education learning is based on physical activity, teachers must be able to internalize measurement and assessment activities on students' analytical thinking skills through the availability of practical and credible instruments.

The study results ensure that the innovative instrument meets the criteria of validity and reliability. Overall, this instrument is suitable for testing the analytical thinking skills of elementary school students in the game of throw–and–catch. Apart from the practical aspect of the question in terms of numbers, this instrument also has questions that are limited to throwing and catching game activities. For example, throwing, catching, and how students pass and catch the ball in a position that is not ideal for stimulating students to explore higher–order thinking skills.

In addition, this instrument also has a "quite difficult" level of difficulty and has a "sufficient" discriminatory power. Thus, in learning motion (physical education), the teacher can assess students' analytical thinking skills, especially throwing and catching game activities. We also suggest that other research can develop analytical thinking skills instruments by using other indicators and for other materials or physical activities

so that in learning physical education, teachers do not ignore students' higher–order thinking skills as a student's need for self–development in order to answer the demands of the world of work in the future.

ACKNOWLEDMENT

The research team would like to thank Dr. Ir. Ayub U. I. Meko, M.Si., as the Rector of Universitas Kristen Artha Wacana, who sponsored research activities in the University's Leading Research Scheme through contract number: 65a/LP–UKAW/E.19/X.2021. We also thank the Center for Education Financing Services and the Scholarship of Educational Fund Management Institutions of the Republic of Indonesia for sponsoring the publication of research results.

REFERENCES

Abosalem, Y. (2016). Assessment techniques and students' higher–order thinking skills. *International Journal of Secondary Education*, *4*(1), 1–11. Retrieved from http://www.sciencepublishinggroup.com/journal/paperinfo?journalid=193&doi=10.11648/j.ijsedu.20160401.11

Ad'hiya, E., & Laksono, E. W. (2018). Development and validation of an integrated assessment instrument to assess students' analytical thinking skills in chemical literacy. *International Journal of Instruction*, *11*(4), 241–256. https://doi.org/10.12973/iji.2018.11416a

Aiken, L. R. (1985). Three coefficients for analyzing the reliability, and validity of ratings. *Educational and Psychological Measurement*, 45, 131–142. https://doi.org/10.1177/0013164485451012

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Ma Ver, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A taxonomy for learning, teaching, and assessing: Revision of Bloom's taxonomy of educational objectives*. Boston: Addison Wesley Longman, Inc.

Anggraeni, D. M., & Sole, F. B. (2020). Analysis of science teachers' understanding of high order thinking skills (HOTS) and their implementation in learning. *Jurnal Penelitian Pendidikan IPA*, 6(2), 209–214. https://doi.org/10.29303/jppipa.v6i2.411

Bierera, E., & Muchlis. (2021). Pengembangan LKPD berbasis PBL untuk melatihkan keterampilan berpikir analitis pada materi elektrolit dan nonelektrolit. *UNESA Journal of Chemical Education*, *10*(2), 149–158. Retrieved from https://ejournal.unesa.ac.id/index.php/journal-of-chemical-education/article/view/38804

Ceniza, J. C., & Cereno, D. C. (2012). Development of mathematic diagnostic test for DORSHS. Retrieved 29 November, 2021 from http://www.doscst.edu.ph/index.ph[/academics/graduateschool/publication/category/5volum-1-issue-12012?

Darmawan, P. (2020). Students' analytical thinking in solving problems of polygon areas. *Kontinu: Jurnal Penelitian Didaktik Matematika*, *4*(1), 17–32. http://dx.doi.org/10.30659/kontinu.4.1.17-32

Fitriani, D., Suryana, Y., & Hamdu, G. (2018). Pengembangan instrumen tes higher–order thinking skill pada pembelajaran tematik berbasis outdoor learning di SD. *Pedadidaktika: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, *5*(1), 252–262. Retrieved from https://ejournal.upi.edu/index.php/pedadidaktika/article/view/7348/0

Fitriani, Fadly, W., & Faizah, U. N. (2021). Analisis keterampilan berpikir analitis siswa pada tema pewarisan sifat. *Jurnal Tadris IPA Indonesia*, *1*(1), 55–67. Retrieved from https://ejournal.iainponorogo.ac.id/index.php/jtii/article/view/64

Fitriyana, N., Marfuatun, & Priyambodo, E. (2019). The profile of students' analytical thinking skills on chemistry systemic learning approach. *Scientiae Educatia: Jurnal Pendidikan Sains*, *8*(2), 207–219. https://doi.org/10.24235/sc.educatia.v8i2.5272

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis*. New Jersey: Pearson University Press.

Hasyim, F. (2018). Mengukur keterampilan berpikir analitis dan keterampilan proses sains mahasiswa calon guru fisika STKIP Al Hikmah Surabaya. *Jurnal Pendidikan IPA Veteran*, *2*(1), 80–89. https://doi.org/10.31331/jipva.v2i1.591

Ho, W. K. Y., Ahmed, M. D., & Kukurova, K. (2021). Development and validation of an instrument to assess quality physical education. *Educational Assessment & Evaluation*, *8*, 1864082. https://doi.org/10.1080/2331186X.2020.1864082

Huda, H. B., & Rohaeti, E. (2020). Research–oriented collaborative inquiry learning (REORCILEA) model: Improvement of students' analytical thinking ability in high school chemistry learning. *Advances in Social Science, Education and Humanities Research, 551*, 248–252. https://doi.org/10.2991/assehr.k.210326.034

Ikmah, I., Nugroho, S. E., & Sutikno. (2018). Analytical thinking skills of vocational students in circular motion cases. *Physics Communication*, *2*(2), 141–150. https://doi.org/10.15294/physcomm.v2i2.14682

Irawati, T. N., & Mahmudah, M. (2018). Pengembangan instrumen keterampilan berpikir analisis siswa smp dalam menyelesaikan soal pemecahan masalah matematika. *Kadikma*, *9*(2), 1–11. https://doi.org/10.19184/kdma.v9i2.8529

Karim, S. A., Sudiro, S., & Sakinah, S. (2021). Utilizing test items analysis to examine the level of difficulty and discriminating power in a teacher–made test. *EduLite: Journal of English Education, Literature, and Culture*, *6*(2), 256–269. http://dx.doi.org/10.30659/e.6.2.256-269

Khairunnisa, N., Sutjihati, S., & Retnowati, R. (2021). Analisis sequential exploratory keterampilan berpikir analisis dalam pembelajaran biologi di SMA Negeri 1 Ciawi. *EduBiologia: Biological Science and Education Journal*, *1*(1), 68–77. http://dx.doi.org/10.30998/edubiologia.v1i1.8229

Laksono, E. W., Rohaeti, E., Suyanta, & Irwanto. (2017). Instrumen penilaian keterampilan berpikir analitis dan keterampilan proses sains kimia. *Jurnal Kependidikan*, *1*(1), 100–110. https://doi.org/10.21831/jk.v1i1.8205

Li, Z., Xia, S., Wu, X., & Chen, Z. (2018). Analytical thinking style leads to more utilitarian moral judgments: An exploration with a process–dissociation approach. *Personality and Individual Differences*, 131(1), 180–184. https://doi.org/10.1016/j.paid.2018.04.046

Maghfiroh, U., & Sugianto. (2011). Penerapan pembelajaran fisika bervisi sets untuk meningkatkan keterampilan berpikir analitis peserta didik kelas X. *Jurnal Pendidikan Fisik Indonesia*, *7*(1), 6–12. https://doi.org/10.15294/jpfi.v7i1.1061

Maksum, A. (2012). *Metodologi penelitian dalam olahraga*. Surabaya: Unesa University Press.

Mayarni, M., & Nopiyanti, E. (2021). Critical and analytical thinking skill in ecology learning: A correlational study. *Jurnal Pendidikan Biologi Indonesia*, *7*(1), 63–70. https://doi.org/10.22219/jpbi.v7i1.13926

Montaku, S. (2011). Results of analytical thinking skills training through students in system analysis and design course. In Proceedings of the IETEC'11 Conference, Kuala Lumpur, Malaysia.

Oriondo, L. L., & Antonio, E. M. D. (1998). *Evaluating educational outcomes: Test, measurement and evaluation*. Manila: Rex Book Store.

Polit, D. F., & Beck, C. T. (2008). Nursing research: Generating and assessing evidence for nursing practice. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Prawita, W., Prayitno, B. A., & Sugiyarto (2019). Students' profile about analytical *thinking skill on respiratory system subject material. Journal of Physics: Conf. Series, 1157*(2), 1–5. Retrieved from https://iopscience.iop.org/article/10.1088/1742-6596/1157/2/022078/meta

Rahayu, P. S., Handoyo, B., & Purwito, H. (2021). Penerapan pendekatan saintifik dilengkapi LKS pada mata pelajaran Geografi untuk meningkatkan keterampilan berpikir analitis siswa kelas XI IPS 2 SMAN 8 Malang. *Jurnal Integrasi dan Harmoni Inovatif Ilmu–Ilmu Sosial*, *1*(3), 336–349. https://doi.org/10.17977/um063v1i3p336-349

Raiyn, J. (2016). The role of visual learning in improving students' high–order thinking skills. *Journal of Education and Practice*, *7*(24), 115–121.

Ramdiah, S., Mayasari, S., Husamah, & Fauzi, A. (2018). The effect of TPS and PBL learning models to the analytical ability of students in biology classroom. *Asia–Pacific Forum on Science Learning and Teaching*, *19*(2), 1–15. Retrieved from https://www.eduhk.hk/apfslt/download/v19_issue2_files/ramdiah.pdf

Ramirez, R. P. B., & Ganaden, M. S. (2008). Creative activities and students' higher order thinking skills. *Education Quarterly*, *66*(1), 22–33. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.824.9279&rep=rep1&type=pdf

Reckase, M. D. (1979). Unifactor laten trait models applied to multifactor test: Result and implications. *Journal of Educational Statistics*, *4*(3), 207–230. https://doi.org/10.3102/10769986004003207

Sadhu, S., & Laksono, W. (2018). Development and validation of an integrated assessment for measucritical thinking and chemical literacy in chemical equilibrium. *International Journal of Instruction*, *11*(3), 557–572. https://doi.org/10.12973/iji.2018.11338a

Sharma, L. R. (2021). Analysis of difficulty index, discrimination index and distractor efficiency of multiple choice questions of speech sounds of English. *International Research Journal of MMC*, *2*(1), 15–28. https://doi.org/10.3126/irjmmc.v2i1.35126

Smits, N., Cuijpers, P., & van Straten, A. (2011). Applying computerized adaptive testing to the CES–D scale: A simulation study. *Pyschiatry Research*, *188*(1), 147–155. https://doi.org/10.1016/j.psychres.2010.12.001

Sudibyo, E., Jatmiko, B., & Widodo, W. (2016). The effectiveness of CBL model to improve analytical thinking skills the students of sport science. *International Education Studies*, *9*(4), 195–203. https://doi.org/10.5539/ies.v9n4p195

Sudijono, A. (2011). *Evaluasi pendidikan*. Jakarta: Raja Grafindo Persada

Sumaryati, S., Joyoatmojo, S., Wiryawan, S. A., & Suryani N. (2020). Potential of E–CoPAL strategy to improve analytical problem solving and teamwork skills in accounting education. *International Journal of Instruction*, *13*(2), 721–732. https://doi.org/10.29333/iji.2020.13249a

Sunyoto, D. (2011). *Analisis regresi dan uji hipotesis*. Yogyakarta: CAPS.

Thaneerananon, T., Triampo, W., & Nokkaew, A. (2016). Development of a test to evaluate students' analytical thinking based on fact versus opinion differentiation. *International Journal of Instruction*, *9*(2), 123–138. https://doi.org/10.12973/iji.2016.929a

Thorndike, R. L., & Hagen, E. (1977). *Measurement and evaluation in psychology and education*. New York: John Wiley and Sons, Inc.

Tosyali, F., & Aktas, B. (2021). Does training analytical thinking decrease superstitious beliefs? Relationship between analytical thinking, intrinsic religiosity, and superstitious beliefs. *Personality and Individual Differences*, *183*. 111122. https://doi.org/10.1016/j.paid.2021.111122

Wiyarsi, A., Fachriyah, A. R., Supriadi, D., & Damanhuri, M. I. M (2019. A test of analytical thinking and chemical representation ability on 'rate of reaction' topic. *Cakrawala Pendidikan*, *38*(2), 228–242. https://doi.org/10.21831/cp.v38i2.23062

Wu, Q., Zhang, Z., Song, Y., Zhang, Y., Zhang, Y., Zhang, F., LI, R., & Miao, D. (2013). The development of mathematical test based on item response theory. *International Journal of Advancements in Computing Technology*, *5*(10), 209–216.

Yulina, I. K., Permanasari, A., Hernani, H., & Setiawan, W. (2019). Analytical thinking skill profile and perception of pre service chemistry teachers in analytical chemistry learning. *Journal of Physics: Conf. Series*, *1157*(4), 1–7. Retrieved from https://iopscience.iop.org/article/10.1088/1742-6596/1157/4/042046

Yuwono, G. R., Sunarno, W., & Aminah, N. S. (2020). Pengaruh keterampilan berpikir analitis pada pembelajaran berbasis masalah (PBL) terhadap hasil belajar ranah pengetahuan. *Edusains*, *12*(1), 106–112. https://doi.org/10.15408/es.v12i1.11659