Research Article

Development of Natural and Social Science Teaching Materials Using the Project-Based Learning Model Oriented Toward Tri Kaya Parisudha to Enhance Learning Motivation and Critical Thinking Skills of Elementary School Students

I Made Astra Winaya[1]*, I Wayan Suastra[2], I Nengah Suastika[3], I Wayan Kertih[4]

1,2,3,4Ganesha University of Education, Bali, Indonesia
Corresponding Author: astra@undiksha.ac.id, iwsuastra@undiksha.ac.id, nengah.suastika@undiksha.ac.id, wayan.kertih@undiksha.ac.id

Abstract

Background: The application of project-based learning methods in Natural Sciences and Social Sciences subjects has not integrated local wisdom in building student character, so the method needs to be oriented towards Tri Kaya Parisudha to increase learning motivation and critical thinking skills of elementary school students. Aim: This research aims to develop Natural and Social Science teaching materials using the Project Based Learning (PjBL) model oriented toward Tri Kaya Parisudha to enhance the learning motivation and critical thinking skills of elementary school students Methods: The research uses a research and development method with the 4D model, which includes Define, Design, Develop, and Disseminate. The research subjects consisted of fourth-grade elementary school students in Marga District, Tabanan Regency, with an experimental group of 36 students and a control group of 40 students. Data collection used a learning motivation questionnaire, observation sheets, and critical thinking skills tests. The validity of the products and instruments was analysed using the Aiken's V index, while the effectiveness was analysed through ANOVA and MANOVA. Results: The research results showed that the teaching materials had very high validity in the aspects of language, content, and media, with Aiken's V values of 0.883, 0.900, and 0.917, respectively. The practicality test obtained a percentage of 76.35%, categorized as quite practical. The critical thinking skills of the experimental group averaged 85.42, higher than the control group at 55.88. Conclusion: Material on learning Natural Sciences and Social Sciences based on the Project-Based Learning Model oriented to Tri Kaya Parisudha is suitable and potentially effective in supporting contextual learning and the development of critical thinking skills in elementary school students.

Keywords: teaching materials, Tri Kaya Parisudha

1. Introduction

At the elementary school level, critical thinking skills are related to the learning motivation possessed by an individual. Learning motivation is the overall driving force within students that-initiates learning activities, ensures the continuity of learning activities, and provides direction to learning activities, so that the desired activities are achieved. Students with high learning motivation are likely to achieve high learning outcomes as well, meaning that the higher the motivation, the more intense the effort and endeavor put in, and thus the higher their critical thinking ability (Dumont & Ready, 2020; Hu et al., 2022).

Based on the preliminary study conducted, through observation, interviews, document studies, and literature reviews, related to learning in elementary schools in Tabanan Regency, the quality of learning is still low. Interviews, observations, and documentation studies were conducted in the Ki Hajar Dewantara Cluster, Marga District, Tabanan Regency, namely State Elementary School 1 Selanbawak, State Elementary School 2 Selanbawak, and State Elementary School 3 Selanbawak. The results of the observations, interviews, document studies, and literature reviews can be described as follows. The results of interviews and observations with several teachers at the school indicate that learning in elementary schools tends to prioritize cognitive aspects, neglecting the affective and psychomotor aspects of students. Learning that overly focuses on cognitive aspects anpd neglects affective and psychomotor aspects causes students to lose intrinsic motivation to learn and opportunities to develop critical thinking skills. Learning is still dominated by activities such as memorization, answering test questions, and pursuing academic achievements, which causes students to feel burdened and less emotionally engaged in the learning process. Moreover, overly structured and test-oriented learning makes students less accustomed to discussing, collaborating, and exploring new ideas. In

classroom learning, students should be given the opportunity to develop an open attitude toward new ideas, the courage to question established norms, and empathy in understanding other people's perspectives. With a more holistic approach, students' motivation to learn and critical thinking skills can develop more optimally (Hallsen, 2021; Nestler et al., 2022; Simpson, 2021).

The results of the observation show that although the curriculum is oriented toward competency mastery, the focus of learning still tends to be oriented toward material mastery (subject-oriented). When given lesson material that requires memorization, it is indeed observed that students are quite proficient or able to memorize the material provided by the teacher and can also explain the material fluently. However, it is different when given an assignment to analyze the material; students tend to explain it not with their own thoughts but with sentences that are almost identical to those in the book they used. Students are unable to develop the content of the material using their own sentences. At the end of the lesson, students are also unable to summarize each subject matter that has been studied. Such a learning process indicates that there are problems in the education system that lead to students' low critical thinking skills, which in turn results in poor learning outcomes. Critical thinking is a mental and intellectual skill that leads to the ability to reason, infer, correlate, and analyze. The ability to think critically is very important for students' life skills in problem-solving (Kuijk et al., 2021; Ross & Morrison, 2021).

Based on a document study of the teaching materials, it was revealed that the device has not yet integrated the local wisdom of the community. This is reinforced by field observation results, where the learning does not connect the material with the students' real-life experiences. As a result, the learning process fails to attract students' interest and lacks significance. Fifth, based on the results of the literature study, learning at the elementary school level in Bali does not optimize the use of local culture. This is in line with observations in several elementary schools in Bali, which show that the integration of local culture into learning is still rarely done. The educational process will never be separate from local wisdom. through the incorporation of real portraits of the socio-cultural dimension into the classroom, it is hoped that students will feel they are learning within the reality of their daily lives, thus avoiding a shock-learning situation (Dyson et al., 2022; Tackie, 2022).

Seeing the above facts, the need for teaching materials by exploring Balinese local wisdom is something that must be done. The model of learning based on local wisdom that is needed is a model that can optimize students' learning motivation and critical thinking skills. This research will focus on the development of Natural and Social Science teaching materials using the *Tri Kaya Parisudha*-based project-based learning model to enhance elementary school students' learning motivation and critical thinking skills. The *Tri Kaya Parisudha*-based project-based learning model, rooted in Balinese local wisdom, represents the character value of self-cultivation and plays an important role in shaping students' mentality. Students will become accustomed to prioritizing not only grades as the result of an assessment but also the quality of what they have gained from the learning process. The development of Natural and Social Science teaching materials using the *Tri Kaya Parisudha*-based project-based learning model to enhance elementary school students' learning motivation and critical thinking skills becomes a basis for critical thinking and solutions, which are important elements, as well as a commitment of a learner in improving non-ideal conditions. We can read that resilience as a character will lead students to understand its essence as public citizens (Ivemark & Ambrose, 2021; Slavin et al., 2021). Transformation or improvement is not achieved through stagnation, but through the process of adaptation and resilience in all conditions. Based on the background, problem identification, and problem limitations that have been outlined, the formulation of the problem in this research is as follows: a) how practical is the use of Natural and Social Science teaching materials with the Project Based Learning model based on *Tri Kaya Parisudha*? ; b) how effective is the Natural and Social Sciences teaching materials with the Project Based Learning model based on *Tri Kaya Parisudha* in increasing the learning motivation and critical thinking skills of elementary school students?

2. Materials And Methods

This research is a development study. The chosen development model refers to the 4D model. The 4D development model developed by Thiagarajan is one of the systematic approaches in the research and development of learning devices. This model consists of four main stages: Define, Design, Develop, and Disseminate. In relation to this research, the 4D model is highly relevant. In the Define stage, the values of *Tri Kaya Parisudha* (thinking good, speaking good, doing good) can be integrated from the analysis of needs and learning objectives, so that the teaching materials not

only focus on the content of Natural and Social Sciences but also build the moral and character dimensions of students.

The Design stage allows for the design of project-based learning activities that align with the local context of Balinese Hindu culture, such as a school environmental management project that instills *Tri Kaya Parisudha*. Next, the Develop stage provides space for validation by experts in Natural and Social Sciences, Hindu religious education, and PjBL learning, ensuring that the resulting products meet academic, pedagogical, and religious standards. Finally, the Disseminate stage provides an opportunity to distribute the teaching materials to other elementary schools, while also making an innovative contribution to the development of education based on local wisdom. Thus, the 4D model is not only methodologically relevant for the development of teaching materials but also aligns with the research objectives to enhance students' learning motivation and critical thinking skills through a contextual, project-based approach rooted in the values of *Tri Kaya Parisudha.*

The population in this study consists of all elementary school students in the Marga District, specifically in Selanbawak Village, during the 2025/2026 academic year. This population includes fourth-grade elementary school students who are studying Natural Sciences and Social Studies. The selection of fourth grade is based on the consideration that at this level, students have adequate basic skills in reading, writing, and arithmetic, and are at the stage of concrete operational cognitive development moving toward formal, making it suitable for the application of project-based learning models that require active involvement and critical thinking.

Teble 1. Research Population

No.	School Name	Number of 4th Grade Students
1	SDN 1 Selanbawak	11
2	SDN 2 Selanbawak	20
3	SDN 3 Selanbawak	16
4	SDN 1 Cau Belayu	22
5	SDN 2 Cau Belayu	18

Then, one class from the elementary school that met the requirements was obtained as the experimental group, and another class as the control group. Based on these criteria, the Experimental Group was then selected, which consisted of Elementary School 2 Selanbawak and Elementary School 3 Selanbawak with a total of 36 people. Subsequently, the Control Group consisted of Elementary School 1 Cau Belayu and Elementary School 2 Cau Belayu with a total of 40 people. Before being used for the main data collection, the learning motivation questionnaire instrument was first tested for validity to ensure that each statement item can accurately and consistently measure the intended aspect. The validity test in this study was conducted through two approaches, namely content validity and empirical (construct) validity.

Content validity aims to assess the extent to which the items in the questionnaire represent all indicators of the learning motivation variable. The assessment is conducted through expert validation, which involves seeking the opinions of supervisors and learning evaluation experts regarding the alignment between the questionnaire statements and the indicators of learning motivation theory. The evaluation includes aspects of readability, language appropriateness for the developmental level of elementary school students, and the relevance of the content to the indicators of learning motivation. Input from experts is used as a basis to revise or refine the questionnaire items before they are used in empirical testing. Empirical validity is conducted to measure the extent to which each questionnaire item correlates with the total score, which reflects the overall construct of learning motivation. This test is conducted by pilot testing the questionnaire on 20–30 respondents who have similar characteristics to the research subjects.

The data from the trial are analyzed using the Pearson Product Moment correlation technique, by comparing the calculated r value against the table r value at a significance level of 5%. Questionnaire items are categorized as valid if the calculated r value > table r value, and conversely considered invalid if the calculated r value ≤ table r value. Invalid items will be revised or eliminated from the instrument to avoid affecting the accuracy of the collected data. By

implementing both forms of validity, the learning motivation questionnaire instrument is expected to have a quality that is suitable for measuring the intended variable accurately and authentically.

In addition to the validity test, the learning motivation questionnaire instrument is also tested for reliability to determine the extent to which the instrument is consistent and stable in measuring the same variable when used at different times or under different conditions. The reliability of the instrument indicates the degree of trust (reliability) in the measurement results. The reliability test is conducted after the instrument is declared valid based on the results of content validity and empirical validity tests. The analysis technique used is Cronbach's Alpha, which is the most common method for measuring reliability for instruments in the form of Likert scales. The calculation of the reliability coefficient is performed with the help of statistical programs (such as SPSS), and the results are compared with the reliability interpretation standards.

If the Cronbach's Alpha value is ≥ 0.60, the instrument is considered to have adequate reliability and is suitable for use in data collection for research. Conversely, if the coefficient is below this minimum threshold, revisions to some items or the entire instrument are necessary. Thus, through this reliability test, it can be ensured that the learning motivation questionnaire provides consistent and trustworthy results in measuring the learning motivation of elementary school students.
The observation instrument used in this study aims to measure the critical thinking skills of elementary school students through real behaviors observed during the learning process. Before being used in the field trial stage, the observation sheet was first tested for validity to ensure that the observed indicators and aspects accurately represent the critical thinking construct. To determine the effectiveness of the developed teaching materials, data analysis methods were subsequently employed. The collected data were tabulated for mean and standard deviation concerning learning motivation and critical thinking skills. The statistical analysis used to test the hypothesis is MANOVA. This study investigates the effect of an independent variable on two dependent variables. Furthermore, the research design can be seen in the table below.

Table 2. Research Data Analysis Design

A1		**A2**	
Y1	**Y2**	**Y1**	**Y2**

The multivariate normality test is conducted to examine the normality of the dependent variables collectively. The multivariate normality referred to is the multivariate normality of the dependent variables within each population (group). The assumption of multivariate normality for dependent variables needs to be conducted because the testing in MANOVA uses Wilk's Lambda, which is distributed as chi-square.

To determine whether there is a sufficiently high relationship between the learning motivation variable and students' critical thinking skills. If there is not a sufficiently high relationship, it means that there are no similar aspects measured in the variable, thus the analysis can continue. The technique used to determine multicollinearity is by using SPSS 25 for Windows with the VIF (variance inflation factors) testing value. If the VIF value is around 1 or close to 1, then there is no multicollinearity problem in the regression model, and if the correlation coefficient between variables is less than 0.5, it means there is no multicollinearity problem.

Hypothesis Testing 1 uses ANOVA through the F variance statistic (between groups). The testing criterion is that if the F value with a significance level of less than 0.05, then H0 is rejected, meaning there is a significant difference in learning motivation between students who participate in learning with the Tri Kaya Parisuda-based Project Based Learning model teaching materials and students who do not receive the treatment.

The statistical analysis used to test the hypothesis of this research employs ANOVA. Hypothesis Testing 2 uses ANOVA

through the F variance statistic (between groups). The testing criterion is that if the F value with a significance level of less than 0.05, then H0 is rejected, meaning there is a significant difference in the critical thinking skills of students who follow learning using the Tri Kaya Parisuda-based Project Based Learning model teaching materials compared to students who follow learning without the developed teaching materials.

The results of the ANOVA test indicate the presence or absence of a global difference, without providing detailed information about which groups differ significantly. Therefore, a post-hoc test is needed, one of which is the Tukey's Honest Significant Difference (Tukey HSD) test. Here is the formula for the Tukey Test that will be used.

$$Q = \frac{|\bar{X}_a - \bar{X}_b|}{\sqrt{\frac{MSE}{n}}}$$

The use of the Tukey HSD test in this study aims to identify pairs of groups (for example, between the experimental and control groups, or among variations of treatment) that differ significantly in terms of learning motivation or critical thinking skills and to provide a more specific interpretation of the effectiveness of PjBL-based teaching materials with the *Tri Kaya Parisudha* value compared to other learning approaches. The results of this study indicate that the Tukey test is capable of more precisely identifying groups of students that differ significantly in their learning outcomes. Thus, in the context of research on the development of Natural and Social Science teaching materials based on Project Based Learning with the *Tri Kaya Parisudha* value, the Tukey test serves to clarify the effectiveness of the intervention by explicitly showing which groups have significantly improved their learning motivation and critical thinking skills.

3. Results

One of the main advantages of the developed teaching materials is the integration of Balinese local wisdom values, namely *Tri Kaya Parisudha*, into every stage of the learning process. The integration is not only realized in the presentation of the material but also implemented through various learning activities carried out by the students. The value of *Manacika* is applied through activities such as problem analysis, critical thinking, and decision-making based on facts and data obtained during the project implementation. The value of *Wacika* is realized through group discussion activities, project result presentations, interviews with resource persons, and the habituation of using polite language and respecting others' opinions. The value of *Kayika* is implemented through field observation activities, group collaboration, responsible project completion, and various real activities that reflect positive behavior. The integration of these three values makes the teaching materials not only oriented toward achieving academic competence but also supports the character formation of students in accordance with local cultural values.

Content validity testing is conducted to determine the degree of alignment of each item in the instrument with the measured indicators before the instrument is used in the research. Content validity is a form of validity that aims to ensure that each item of the instrument adequately represents the construct to be measured (Azwar, 2019). In this study, the content validity of the linguistic aspect instrument was conducted through expert judgment by Dr. I Kadek Suartama, S.Pd., M.Pd. and Prof. Dr. I Wayan Suastra, M.Pd. as validators. The assessment was conducted using a four-point scale, namely score 1 (not relevant), score 2 (less relevant), score 3 (relevant), and score 4 (very relevant).

The results of the Aiken's V index calculation for each item of the linguistic aspect instrument are presented in the following table.

Table 3. The Aiken's V index calculation results for linguistic aspects

Item Number	Validator 1	Validator 2	Σs	Aiken's V	Criteria
1	4	4	6	1,000	Very Valid
2	4	3	5	0,833	Very Valid
3	4	4	6	1,000	Very Valid
4	4	3	5	0,833	Very Valid

5	3	4	5	0,833	Very Valid
6	4	3	5	0,833	Very Valid
7	4	4	6	1,000	Very Valid
8	4	3	5	0,833	Very Valid
9	3	3	4	0,667	Valid
10	4	4	6	1,000	Very Valid

Based on the table, it is known that the Aiken's V Index value on the linguistic aspect instrument is in the range of 0.667–1.000. Four items received a value of 1.000, five items received a value of 0.833, and one item received a value of 0.667. The overall average value is 0.883, indicating that the instrument has a very high level of content validity. According to Aiken (1985), the closer the value is to 1.00, the higher the level of agreement among experts regarding the relevance of an item in the instrument. Furthermore, Azwar (2019) explains that a high content validity coefficient indicates that the instrument items have represented the measured indicators, making them suitable for use in the research data collection process. Thus, the analysis results indicate that all items of the linguistic aspect instrument have met the content validity criteria and can be used in the next stage of the research.

Nevertheless, item number 9 received an Aiken's V value of 0.667, which is the lowest value compared to the other items. This value still falls within the valid category, so the item can be retained. However, several editorial improvement suggestions provided by the validators were still used as a basis for revision to enhance the clarity of language use, sentence accuracy, and ease of understanding of the instrument by respondents. This revision process was carried out as part of the quality enhancement of the instrument before it is implemented in field trials. In addition to linguistic aspects, content validity is also conducted on the instrument's material/content aspect. This testing aims to determine the degree of alignment between the substance of the material in the instrument and the formulated research indicators. Content validity on the material/content aspect was conducted through expert judgment by Prof. Dr. I Wayan Suastra, M.Pd. and Dr. I Kadek Suartama, S.Pd., M.Pd. using a four-point rating scale, namely score 1 (not relevant), score 2 (less relevant), score 3 (relevant), and score 4 (very relevant).

Content validity analysis uses Aiken's V Index as proposed by Aiken (1985). This method is used to measure the level of agreement among validators regarding the relevance of each item in the instrument. The higher the value of the Aiken's V coefficient, the higher the level of agreement among experts that an item in the instrument is in accordance with the measured indicator. The results of the Aiken's V Index calculations for each item of the content aspect instrument are presented in the table below.

Table 4. The Content Validity Test Results for Material/Content Aspect Instrument

Item Number	Validator I	Validator II	Σs	Aiken's V	Criteria
1	4	4	6	1,000	Very Valid
2	4	4	6	1,000	Very Valid
3	4	3	5	0,833	Very Valid
4	4	4	6	1,000	Very Valid
5	4	3	5	0,833	Very Valid
6	4	4	6	1,000	Very Valid
7	4	3	5	0,833	Very Valid
8	4	4	6	1,000	Very Valid

9	3	4	5	0,833	Very Valid
10	4	3	5	0,833	Very Valid

Based on the analysis results in the Table, the Aiken's V Index values range from 0.833 to 1.000. Five items received a score of 1.000, while the other five items received a score of 0.833. There are no items with a value below 0.80, so all items of the content aspect instrument fall into the very valid category. The average Aiken's V Index score of 0.900 indicates that the instrument has a very high content validity level. The results indicate a high level of agreement between the two validators regarding the content suitability, material coverage, concept accuracy, and the relevance of each item in the instrument to the measured indicators. According to Aiken (1985), a validity coefficient approaching 1.00 indicates that the experts provide very consistent assessments of the relevance of each item in the instrument. In line with that opinion, Azwar (2019) states that high content validity indicates that the instrument has been able to comprehensively represent the domain of the construct to be measured.

Based on the results, it can be concluded that all items in the instrument for the material/content aspect have met the content validity criteria and are suitable for use as a research instrument. Although some items received a score of 0.833, the validators still provided several suggestions regarding the refinement of wording, accuracy of terminology, and adjustment of examples to better align with the characteristics of the students. The input was used as a basis for making minor revisions so that the quality of the instrument became more optimal before being used in the field trial stage. The content validity test on the media aspect is conducted to determine the degree of suitability of the instrument in measuring the quality of the developed learning media. The media aspect includes indicators of visual appearance, interface design, readability, ease of use, illustration quality, layout consistency, color combination, typography, and media appeal. The assessment was conducted through expert judgment by Prof. Dr. I Wayan Suastra, M.Pd. and Dr. I Kadek Suartama, S.Pd., M.Pd. using a four-point rating scale, namely score 1 (not relevant), score 2 (less relevant), score 3 (relevant), and score 4 (very relevant).

Content validity analysis is conducted using Aiken's V Index as developed by Aiken (1985). This index is used to determine the level of agreement among experts regarding the relevance of each item in the instrument used to evaluate the quality of learning media. The higher the Aiken's V coefficient value, the higher the degree of alignment of each item in the instrument with the measured construct. The results of the Aiken's V Index calculation on the media aspect instrument are presented in the table.

Table 5. The Content Validity Test Results for Media Aspect Instruments

Item Number	Validator I	Validator II	Σs	Aiken's V	Criteria
1	4	4	6	1,000	Very Valid
2	4	4	6	1,000	Very Valid
3	4	4	6	1,000	Very Valid
4	4	3	5	0,833	Very Valid
5	3	4	5	0,833	Very Valid
6	4	4	6	1,000	Very Valid
7	3	3	4	0,667	Valid
8	4	4	6	1,000	Very Valid
9	4	4	6	1,000	Very Valid

10	4	3	5	0,833	Very Valid

Based on the analysis results in the Table, the Aiken's V Index values range from 0.667 to 1.000. Six items received a value of 1.000, three items received a value of 0.833, and one item received a value of 0.667. Overall, the average Aiken's V Index value is 0.917, which falls into the very valid category. The value indicates that both validators have a very high level of agreement on the quality of each indicator in the media aspect. The majority of items received the maximum score, which indicates that the media appearance, visual design, layout, color usage, typography, illustrations, and ease of use are in accordance with the principles of good educational media development. According to Aiken (1985), a validity coefficient approaching 1.00 indicates a very high level of agreement among validators regarding the relevance of each item in the instrument. This opinion is supported by Azwar (2019), who states that instruments with a high content validity coefficient are capable of accurately representing the measured construct, making them suitable for use in the research data collection process.

Nevertheless, item number 7 received a score of 0.667, which is the lowest score in the media aspect. The score still meets the validity criteria, so the item is retained. The validator provided several suggestions for improvement, particularly regarding enhancing design consistency, layout arrangement, and optimizing visual elements to make the media more attractive, communicative, and user-friendly for learners. All of these inputs were accommodated through minor revisions, making the quality of the media more optimal before being implemented in the field trial stage. Based on the results of the content validity test, it can be concluded that all items of the media aspect instrument meet the content validity criteria, ranging from valid to very valid. Thus, the media aspect instrument is deemed suitable for use as a tool to evaluate the quality of the learning media developed in this research.

If compared among aspects, the media aspect received the highest validity score (0.917), followed by the material/content aspect (0.900) and the linguistic aspect (0.883). The difference is relatively small, indicating that the three aspects have balanced quality and support each other. The high validity score in the media aspect indicates that the developed product has successfully integrated learning design principles, while the high validity in the content and linguistic aspects shows that the product's content aligns with the learning objectives and is presented in a communicative and easily understandable language.

Overall, the expert validation results prove that the instruments and products developed have met the content validity requirements. These results reinforce Aiken's (1985) opinion that the higher the Aiken's V coefficient value, the higher the level of agreement among experts regarding the relevance of an instrument. Furthermore, the results of this study are also in line with Azwar's (2019) opinion, which states that high content validity indicates that each item of the instrument has adequately represented the construct domain, making it suitable for use as a data collection tool. Thus, the research instrument is deemed suitable for use in the field trial stage after minor revisions based on the validators' feedback.

The findings of this study are also supported by previous research results that show that the expert validation process is an important stage in development research to ensure product quality before implementation. Sugiyono (2022) states that expert validation aims to ensure the alignment of content, language, and presentation of the product with the development objectives so that the resulting product has a high level of feasibility. Furthermore, Thiagarajan et al. (1974) in their Four-D (4-D) development model emphasize that the development stage must begin with expert validation to obtain feedback before the product is tested on users. Therefore, the validation results showing a very valid category in this study serve as evidence that the product has met the feasibility standards in terms of language, content, and media, making it suitable to proceed to the practicality and effectiveness testing stages.

4. Discussion

1. The Practicality of Using Natural and Social Science Teaching Materials with the *Tri Kaya Parisudha*-Based Project Based Learning Model

Practicality testing is one of the evaluation stages in research and development (Research and Development) aimed at determining the ease of use of a product based on the direct experience of the target users. According to Nieveen (1999), an educational product is considered of high quality if it meets three main criteria: valid, practical, and

effective. The aspect of practicality indicates the extent to which the product is easy to use, easy to understand, and can be applied in the learning process according to the established objectives. In line with that, Plomp (2013) explains that the practicality of a product is determined based on user assessments of the ease of implementing the product in actual learning conditions.

The practicality test in this study was conducted on 36 fourth-grade students, consisting of 20 students from Sekolah Dasar Negeri 2 Selanbawak and 16 students from Sekolah Dasar Negeri 3 Selanbawak. Data collection was conducted using a student response questionnaire consisting of 20 statements with a four-point Likert scale, namely score 1 (strongly disagree), score 2 (disagree), score 3 (agree), and score 4 (strongly agree). In development research, the empirical validity analysis of user response questionnaires is part of formative evaluation aimed at ensuring that the instruments used to measure practicality meet the quality requirements of research instruments (Fenwick & Edwards, 2016; Robinson et al., 2014).

The validity analysis of the items was conducted using IBM SPSS Statistics with the help of Corrected Item-Total Correlation (CITC). According to Azwar (2021), an item is considered valid if it has a Corrected Item-Total Correlation value ≥ 0.30. Based on the analysis results, the Corrected Item-Total Correlation values range from 0.510 to 0.907. All statement items have values above the minimum threshold of 0.30, so all 20 practicality questionnaire items are declared valid. The results indicate that each item is capable of representing the construct of product practicality and contributes well to the total questionnaire score. The highest correlation value is found in item P13 at 0.907, while the lowest value is in item P5 at 0.510. Nevertheless, all items still meet the criteria for empirical validity, so there are no items that need to be eliminated or revised. Based on the results of the descriptive analysis of the practicality questionnaire filled out by 36 students, an average score of 61.08 was obtained, with a minimum score of 45 and a maximum score of 77. A standard deviation of 7.64 indicates that there is variation in students' assessments of the product's practicality, but the score distribution remains at a relatively moderate level. This indicates that most students provided relatively consistent evaluations of the product's ease of use (Edovald & Nevill, 2021; Razza et al., 2021).

The findings indicate that the practicality questionnaire instrument has good quality and is suitable for obtaining data on students' responses to the developed product. Thus, the results of the item validity test provide empirical evidence that the instrument is capable of adequately measuring the practicality aspect of the product as part of the formative evaluation in development research. Before determining the practicality category, the actual score obtained from all respondents was first calculated.

The results indicate that the developed product can be used in the learning process and has received positive feedback from the students. However, the practicality test results also indicate that there are still several aspects that can be improved, especially based on the direct experiences of the learners while using the product. This condition aligns with the characteristics of development research, which places the practicality test as part of formative evaluation, namely the process of obtaining feedback from users to refine the product before it is implemented more widely. Thus, the results of the practicality test are not intended to test the effectiveness of the product, but rather to identify the ease of use and obtain input from users as a basis for further revisions. Therefore, if there are still some aspects that are considered not optimal by the learners, these results become important information for the researchers to improve the product before entering the effectiveness testing stage (Beycioglu & Kondakci, 2021; Marianno et al., 2022).

2. The Effectiveness of Natural and Social Science Teaching Materials with the *Tri Kaya Parisudha*-Based Project Based Learning Model Can Improve Elementary School Students' Learning Motivation and Critical Thinking Skills

After conducting a descriptive statistical analysis of students' learning motivation, the same analysis was also performed on the critical thinking skills data of students in the experimental group. The descriptive statistical analysis aims to provide an overview of the characteristics of students' critical thinking skills data after participating in learning using Natural and Social Sciences teaching materials with the Project Based Learning model oriented toward *Tri Kaya Parisudha*. The statistics presented include the number of samples, minimum score, maximum score, mean, median, mode, standard deviation, variance, and score range (Benner, 2021; Hansson & Ohman, 2022).

In the variable of critical thinking skills, data is obtained through multiple-choice test instruments, so descriptive

analysis does not use the Ideal Mean (Mi) and Ideal Standard Deviation (SDi). The calculation of Mi and SDi is generally used in non-test instruments, such as questionnaires with Likert scales, because they have maximum and minimum ideal scores theoretically established. Conversely, in test instruments, analysis is conducted based on the empirical scores obtained by the students, so data description is more accurately presented through mean, median, mode, standard deviation, variance, minimum score, maximum score, and frequency distribution (Jones et al., 2022; Yan & Yang, 2021). This approach aims to provide an overview of the tendency of centralization and dispersion of learning outcomes or abilities measured based on the actual achievements of the students, rather than based on ideal score categories. The results of this analysis are used to determine the tendency of students' critical thinking skills before conducting inferential analysis.

The frequency distribution of students' critical thinking skills in the experimental group shows that most students fall within the score interval of 85–89, with 9 students (25.00%). Next, 7 students (19.44%) fall within the interval of 95–100, while 6 students (16.67%) each fall within the intervals of 75–79 and 90–94. In the interval of 80–84, there are 5 students (13.89%), while the interval of 70–74 has the lowest frequency, with 3 students (8.33%). The distribution shows that students' critical thinking skills scores tend to be concentrated in the high-value intervals. This is supported by an average score of 85.42, which indicates that, in general, the critical thinking skills of students in the experimental group are in the high category (Brown, 2022; Duckworth et al., 2021).
Descriptive statistical analysis was also conducted on the critical thinking skills data of students in the control group to provide an overview of the data characteristics after the learning process without using Natural and Social Sciences teaching materials with the Project Based Learning model oriented toward *Tri Kaya Parisudha*. The descriptive statistics analyzed include the number of samples, minimum score, maximum score, mean, median, mode, standard deviation, variance, and score range. The results of this analysis are used to determine the tendency of critical thinking skills among students in the control group and as a basis for comparison with the experimental group before hypothesis testing is conducted.

It is known that the number of students in the control group is 40. The results of the descriptive statistical analysis show that the students' critical thinking skills scores range from 40 to 70 with a score range of 30. The average (mean) score obtained is 55.88, the median is 57.50, and the mode is 55. Additionally, a standard deviation of 9.02 with a variance of 81.35 was obtained. The average score, which is not far from the median and mode, indicates that the distribution of students' critical thinking skills in the control group is relatively even without any significant differences between students' scores. However, when compared to the experimental group, which has an average of 85.42, the critical thinking skills achievement in the control group is still lower. To obtain a more detailed picture of the distribution of students' critical thinking skills scores in each class interval.

The frequency distribution of students' critical thinking skills in the control group shows that the score distribution is relatively even in the medium to high intervals. The score intervals of 55–59, 60–64, and 65–70 each had the highest frequency, with 8 students (20.00%). Next, the 50–54 interval was occupied by 7 students (17.50%), the 45–49 interval by 5 students (12.50%), while the 40–44 interval had the lowest frequency, with 4 students (10.00%). These results indicate that most students in the control group scored in the middle category with an average score of 55.88. In comparison, the experimental group had an average score of 85.42, which descriptively shows that the critical thinking skills of students in the experimental group were higher than those in the control group. This difference indicates that the application of Natural and Social Sciences teaching materials using the *Tri Kaya Parisudha*-oriented Project Based Learning model has the potential to positively contribute to the improvement of students' critical thinking skills.

5. Conclusion

Based on the research results, it can be concluded that the Natural and Social Sciences teaching materials with the Project Based Learning (PjBL) model oriented toward *Tri Kaya Parisudha* were successfully developed through the 4D model, which includes Define, Design, Develop, and Disseminate. The developed teaching materials have very good quality based on expert validation results in the aspects of language, material/content, and media, with average Aiken's V values of 0.883, 0.900, and 0.917, respectively. The product also received a positive response from the students with a practicality level of 76.35%, which falls into the fairly practical category. The application of teaching materials shows a difference in critical thinking skills achievement between the experimental group and the control group. The experimental group, which used PjBL-based teaching materials oriented toward *Tri Kaya Parisudha*,

achieved an average critical thinking skill score of 85.42, while the control group achieved an average of 55.88. The integration of the values of *Manacika*, *Wacika*, and *Kayika* in project activities also provides a learning context that connects critical thinking skills, collaboration, communication, responsibility, and character building. Thus, the Natural and Social Sciences. learning materials based on PjBL oriented toward *Tri Kaya Parisudha* are suitable for use and have the potential to support contextual, meaningful IPAS learning, as well as being oriented toward the development of critical thinking skills in elementary school students.

References

1. X. Hu, J. C. Ortagus, N. Voorhees, K. Rosinger, and R. Kelchen, "Disparate Impacts of Performance Funding Research Incentives on Research Expenditures and State Appropriations," *AERA Open*, vol. 8, no. 1, pp. 1–24, 2022, doi: 10.1177/23328584211071109.
2. H. Dumont and D. D. Ready, "Do Schools Reduce or Exacerbate Inequality? How the Associations Between Student Achievement and Achievement Growth Influence Our Understanding of the Role of Schooling," *Am. Educ. Res. J.*, vol. 57, no. 2, pp. 728–774, 2020, doi: 10.3102/0002831219868182.
3. A. Simpson, "Benchmarking a Misnomer: A Note on 'Interpreting Effect Sizes in Education Interventions,'" *Educ. Res.*, pp. 1–3, 2021, doi: 10.3102/0013189X20985448.
4. S. Nestler, O. Ludtke, and A. Robitzsch, "Analyzing Longitudinal Social Relations Model Data Using the Social Relations Structural Equation Model," *J. Educ. Behav. Stat.*, vol. 47, no. 2, pp. 231–260, 2022, doi: 10.3102/10769986211056541.
5. S. Hallsen, "The Rise of Supplementary Education in Sweden: Arguments, Thought Styles, and Policy Enactment," *ECNU Rev. Educ.*, vol. 4, no. 3, pp. 476–493, 2021, doi: 10.1177/2096531120952096.
6. S. M. Ross and J. R. Morrison, "Achieving Better Educational Practices Through Research Evidence: A Critical Analysis and Case Illustration of Benefits and Challenges," *ECNU Rev. Educ.*, vol. 4, no. 1, pp. 108–127, 2021, doi: 10.1177/2096531120916742.
7. M. van Kuijk, M. M. Wijnsma, and R. Bosker, "A Systematic Review of Studies Addressing the Implementation of the Evidence-Based Whole-School Reform 'Success for All,'" *ECNU Rev. Educ.*, vol. 4, no. 1, pp. 128–163, 2021, doi: 10.1177/2096531120961521.
8. H. N. Tackie, "(Dis)Connected: Establishing Social Presence and Intimacy in Teacher–Student Relationships During Emergency Remote Learning," *AERA Open*, vol. 8, no. 1, pp. 1–14, 2022, doi: 10.1177/23328584211069525.
9. B. Dyson, Y. Shen, W. Xiong, and L. Dang, "How Cooperative Learning Is Conceptualized and Implemented in Chinese Physical Education: A Systematic Review of Literature," *ECNU Rev. Educ.*, vol. 5, no. 1, pp. 185–206, 2022, doi: 10.1177/20965311211006721.
10. B. Ivemark and A. Ambrose, "Habitus Adaptation and First-Generation University Students' Adjustment to Higher Education: A Life Course Perspective," *Sociol. Educ.*, vol. 94, no. 3, pp. 191–207, 2021, doi: 10.1177/00380407211017060.
11. R. E. Slavin, A. C. K. Cheung, and T. Zhuang, "How Could Evidence-Based Reform Advance Education?," *ECNU Rev. Educ.*, vol. 4, no. 1, pp. 7–24, 2021, doi: 10.1177/2096531120976060.
12. T. Fenwick and R. Edwards, "Exploring The Impact of Digital Technologies on Professional Responsibilities and Education," *Eur. Educ. Res. J.*, vol. 15, no. 1, pp. 117–131, 2016, doi: 10.1177/1474904115608387.
13. T. J. Robinson, L. Fischer, D. Wiley, and J. Hilton, "The Impact of Open Textbooks on Secondary Science Learning Outcomes," *Educ. Res.*, vol. 43, no. 7, pp. 341–351, 2014, doi: 10.3102/0013189X14550275.
14. R. Razza, D. Bergen-Cico, S. Reid, and R. L. Uveges, "The Benefits of Mindfulness for Promoting Resilience Among At-Risk Adolescents: Results From the Inner Strength Teen Program," *ECNU Review of Education*. 2021, doi: 10.1177/2096531120982254.
15. T. Edovald and C. Nevill, "Working Out What Works: The Case of the Education Endowment Foundation in England," *ECNU Rev. Educ.*, vol. 4, no. 1, pp. 46–64, 2021, doi: 10.1177/2096531120913039.
16. B. D. Marianno, A. A. Hemphill, A. P. S. Loures-Elias, L. Garcia, D. Cooper, and E. Coombes, "Power in a Pandemic: Teachers' Unions and Their Responses to School Reopening," *AERA Open*, vol. 8, no. 1, pp. 1–16, 2022, doi: 10.1177/23328584221074337.
17. K. Beycioglu and Y. Kondakci, "Organizational Change in Schools," *ECNU Rev. Educ.*, vol. 4, no. 4, pp. 788–807, 2021, doi: 10.1177/2096531120932177.
18. D. Benner, "On Justice in Pedagogical Contexts," *ECNU Rev. Educ.*, vol. 4, no. 4, pp. 669–685, 2021, doi: 10.1177/2096531120933278.
19. P. Hansson and J. Ohman, "Museum Education and Sustainable Development: A Public Pedagogy," *Eur. Educ. Res. J.*,

vol. 21, no. 3, pp. 469–483, 2022, doi: 10.1177/14749041211056443.
20. B. D. Jones, Y. Miyazaki, M. Li, and S. Biscotte, “Motivational Climate Predicts Student Evaluations of Teaching: Relationships Between Students’ Course Perceptions, Ease of Course, and Evaluations of Teaching,” *AERA Open*, vol. 8. 2022, doi: 10.1177/23328584211073167.
21. S. Yan and Y. Yang, “Education Informatization 2.0 in China: Motivation, Framework, and Vision,” *ECNU Rev. Educ.*, vol. 4, no. 2, pp. 410–428, 2021, doi: 10.1177/2096531120944929.
22. A. L. Duckworth *et al.*, “Students Attending School Remotely Suffer Socially, Emotionally, and Academically,” *Educ. Res.*, vol. 50, no. 7, pp. 479–482, 2021, doi: 10.3102/0013189X211031551.
23. G. T. L. Brown, “Student Conceptions of Assessment: Regulatory Responses to Our Practices,” *ECNU Rev. Educ.*, vol. 5, no. 1, pp. 116–139, 2022, doi: 10.1177/20965311211007869.