International Journal of Instruction
e-ISSN: 1308-1470 ● www.e-iji.net

April 2023 ● Vol.16, No.2
p-ISSN: 1694-609X
pp. 413-436

Article submission code: 20220509001839

Received: 09/05/2022
Revision: 09/10/2022

Accepted: 03/11/2022
OnlineFirst: 12/01/2023

Development of Evaluation Instruments to Measure the Quality of Spatial Problem Based Learning (SPBL): CIPP Framework

Waode Yunia Silviariza
Universitas Negeri Malang, Indonesia, *waode.yunia.2007219@students.um.ac.id*

Sumarmi
Universitas Negeri Malang, Indonesia, *sumarmi.fis@um.ac.id*

Sugeng Utaya
Universitas Negeri Malang, Indonesia, *sugeng.utaya.fis@um.ac.id*

Syamsul Bachri
Universitas Negeri Malang, Indonesia, *syamsul.bachri.fis@um.ac.id*

Budi Handoyo
Universitas Negeri Malang, Indonesia, *budi.handoyo.fis@um.ac.id*

Valid and reliable instruments are an important part of the process of evaluating and correcting the quality of a ‘model of teaching’. Therefore, this study aims to develop an 'evaluation instrument' constructed from the CIPP model to measure the quality of the Spatial Problem Based Learning. This study uses research and development methods; 4D Models. Data collection through interviews and questionnaires. Product validity analysis was obtained from a questionnaire using V'Aikens. While the reliability analysis uses Inter Class Correlation (ICC). Interview data became supporting qualitative data. This research was then reviewed by an expert jury and four practitioners (geography teachers) as evaluators, and twenty-two high school students in geography class as users. Analysis of the validity and reliability of the questionnaire data was carried out with the help of SPSS. The product 'evaluation instrument' has a V'Aikens value of 0.63 from evaluators and 0.78 from users stating that the product 'evaluation instrument' is categorized as “medium”. The ICC value of 0.781 means that the product 'evaluation instrument' is in the “good” category. In addition, the expert states and supports that the product 'evaluation instrument' can measure the quality of the ‘model of teaching’. The results show that the product 'evaluation instrument' constructed by CIPP has a category that is not only valid but also reliable. The 'evaluation instrument' is then used to evaluate the quality of the spatial problem-based learning.

Keywords: CIPP, evaluation instrument, model learning, teaching, learning

Citation: Silviariza, W. Y., Sumarmi., Utaya, S., Bachri, S., & Handoyo, B. (2023). Development of evaluation instruments to measure the quality of spatial problem based learning (SPBL): CIPP framework. *International Journal of Instruction*, *16*(2), 413-436. https://doi.org/10.29333/iji.2023.16223a

INTRODUCTION

A good 'model of teaching' is the product of a long-term investigation (Joyce & Weil, 2003) and continues through the evaluation process (Nieveen, & Folmer, 2013; Dick, et al, 2013). The systematic evaluation process aims to determine the quality of the 'model of teaching'. The evaluation aims to describe, obtain and provide useful information for assessing (Zhang, et al, 2011; Stufflebeam, & Shinkfield, 2012) in this case the 'model of teaching'. Evaluation of the 'model of teaching' can be done in a formative manner. The goal is to account for, validate, and determine the reliability (Yusuf, 2017) that the 'model of teaching' remains of high quality.

Researchers and development should continue to evaluate the shortcomings and constraints of a Spatial Problem-Based Learning (SPBL) model so that it remains of high quality. After going through the development process and empirical studies, there are several obstacles in the implementation process. Empirical studies show that the SPBL model affects the critical thinking skills of students in Geography class (Silviariza, et al, 2021). Although there is an effect, the SPBL model has not been effective. This is evident from the results of the analysis of the N-gain score calculation data that the effectiveness of SPBL is <40% (Hake, R. R, 1999). Another finding is that activities in the SPBL syntax are less systematic and require a long duration of implementation. The effectiveness of the 'model of teaching' is supported by the ideal implementation duration (Wijnia, 2016; Pourshanazari, et al, 2013; Strobel & Barneveld, 2009) and a systematic syntax (Behar-Horenstein & Seabert, 2005). In addition, in the initial development process, experts as model validators stated "The SPBL Model Syntax has not been consistently related". Therefore, it is important to review the quality of the SPBL model further which can be done by formative evaluation.

Formative evaluation of the SPBL model requires instruments. The evaluation instrument must be able to measure and assess the SPBL model comprehensively (Sanjaya, 2015). The instrument must meet certain requirements, provide meaningfully accurate data for its function, and be the only measurement sample. The characteristics of a good instrument are valid, reliable, relevant, representative, practical, descriptive, specific, and proportional (Arifin, 2016; Zhang, et al, 2011).

CIPP provides the widest possible space to assess the context, inputs, processes, and products of the SPBL model. In its context, the SPBL model is based on constructivism with a spatial perspective. The things that become input for a model are of course the syntax, social system, principal of reaction, and support system (Thelen, 1960). Then the important thing to be assessed is how the SPBL model process is implemented and how students respond/attitudes when using the SPBL model. Thus, the study of the product, namely the syntactic sequence, can be evaluated more comprehensively.

Many previous studies have reported on the CIPP model evaluation instrument which is a tool for measuring the quality of a product. Among them are evaluating the Education program (Lippe, & Carter, 2018; Agustina, & Mukhtaruddin, 2019; Iqbal, et al., 2021; Bukit, et al., 2019), book program (Asadi et al., 2016), training program (Umam, & Saripah, 2018), kindergarten education curriculum (Shanawani, 2019; Basaran, et al.,

2021; Aslan, & Uygun, 2019) to higher education (Ebtesam, & Foster, 2019; Tuna, & Başdal, 2021), evaluate students' abilities (Sanusi, et al., 2021) Several previous studies provide references for this study.

Other notes on developing an evaluation instrument based on the CIPP model for portfolio assessment implementation (Kurnia, et al., (2017). In addition, the CIPP model was also developed to evaluate the implementation of Project Assessment in Science Learning (Asfaroh, 2017). Both implemented CIPP to evaluate portfolio assessment in junior high school science learning. The development of the CIPP model instrument in previous studies focused on evaluating product units and or when they were implemented.

In this study, the focus is more on the characteristics and syntax of the SPBL model as a basic reference for compiling evaluation instruments with the CIPP model comprehensively. Thus, this study aims to develop an evaluation instrument product constructed from the Context, Input, Process, and Product (CIPP) Stufflebeam (2003) model. Thus, the instrument is valid and reliable and can be a frame of reference for assessing the 'teaching model' in this case the Spatial Problem Based Learning (SPBL) model with the Formative Evaluation method.

Context, Input, Process Dan Product (CIPP) Model

The Context, Input, Process, and Product (CIPP) evaluation framework is “a comprehensive framework for formative and summative evaluation of a project, personnel, product, organization, and evaluation system" (Stufflebeam, and Coryn, 2014). The CIPP evaluation framework is specially configured to guide a comprehensive and systematic examination of dynamic real-world social or educational projects (Stufflebeam & Shinkfield, 2007). Over the years, the model has been refined and used in a variety of disciplines. Based on a survey by members of the American Society for Training and Development found that the CIPP model is preferred over other evaluation models (Galvin, 1983). In educational settings, the CIPP evaluation model has been used to evaluate various educational projects and entities (Zhang, et al, 2011).

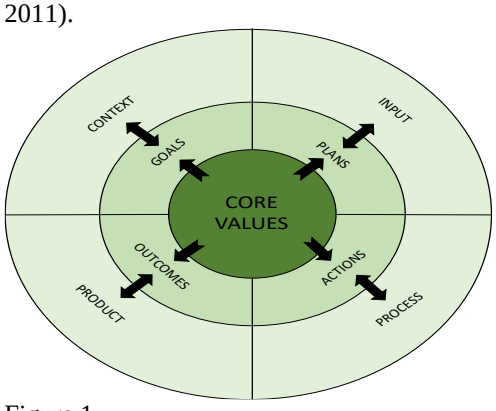

Figure 1
CIPP evaluation model comprehensive framework (Stufflebeam, 2010)

The CIPP evaluation model considers the product as a system, so product evaluation is carried out in detail based on each of its components (Stufflebeam & Shinkfield, 2007; Arikunto & Safruddin, 2014). Proactively, this 4-part evaluation asks "What needs to be done?", "How should it be done?", "What is being done?", "Is it working?". The CIPP evaluation model requires a series of questions to be asked about four different model elements in context, input, process, and product (Arikunto & Jabar, 2009; Tiantong, & Tongchin, 2013). The discipline affects the evaluation instrument.

The CIPP evaluation instrument requires an analysis of what is needed, namely Context, Input, Process, and Product. The CIPP evaluation instrument is designed to systematically guide assessments at the beginning of the project (context and input evaluation), in progress (input and process evaluation), and at the end (product evaluation) (Zhang, et al, 2011). The components of the CIPP model system are as follows:

Context

In this component, evaluators assist researchers in planning decisions, determining product needs, and formulating product goals. In addition, an evaluator also makes decisions on the conditions under which the product will be evaluated and analyzes the needs that have not been met, and identify the reasons behind these needs whether they have not been or have been achieved (Peter, 1992). The purpose of context evaluation is to assess the overall environmental readiness of the project, check whether the existing objectives and priorities are adapted to the needs, and assess whether the proposed objectives are sufficiently responsive to the assessed needs (Stufflebeam, 2003).

Input

One of the evaluations aims to help make decisions, determine sources, alternatives to be taken, what plans and strategies to achieve needs, and how to work procedures will be achieved (Rachmaniar, et al., 2021). The same thing is expressed which states that this evaluation helps to determine the information that will be used to meet the objectives (Stufflebeam, 1985). Examples of sources that influence efforts to achieve goals are the way the teacher teaches, the use of learning media, and the learning environment. Based on the description above, it can be seen that the evaluation of inputs is related to what strategies can be used to achieve needs that have not been or have not been achieved (Stufflebeam, 2000a).

Process

The process evaluation focuses on the implementation decisions that control and manage the product. The process provides feedback on the system or program being examined (Aziz et al., 2018). Furthermore, it is the process of documenting a program and providing feedback and revision on the running program (Stufflebeam, 2010).

Product

Product evaluation is the result that has been achieved from the implementation of a program. Product evaluation aims to measure, interpret, and assess the results and

interpret the benefits, value, and significance, of a product/program (Stufflebeam, 2010). "Product" evaluation activities include gathering descriptions and evaluating outcomes and relating them to objectives and context, input, and information processes, and interpreting their value and benefits (Stufflebeam, 2001). In addition, product evaluation is a process to measure, interpret and assess the extent to which the product can be implemented and achieve the implementation objectives.

CIPP Model for Learning Evaluation Instruments

The construction of the development of the evaluation instrument component of the 'model of teaching' is based on the CIPP model. The CIPP model is a special tool that is useful and simple to help evaluators produce data collection instruments (questions) that are important to be asked in an evaluation process (Hakan & Seval, 2011; Umam, & Saripah, (2018). The following is a framework for evaluating the CIPP-based SPBL model.

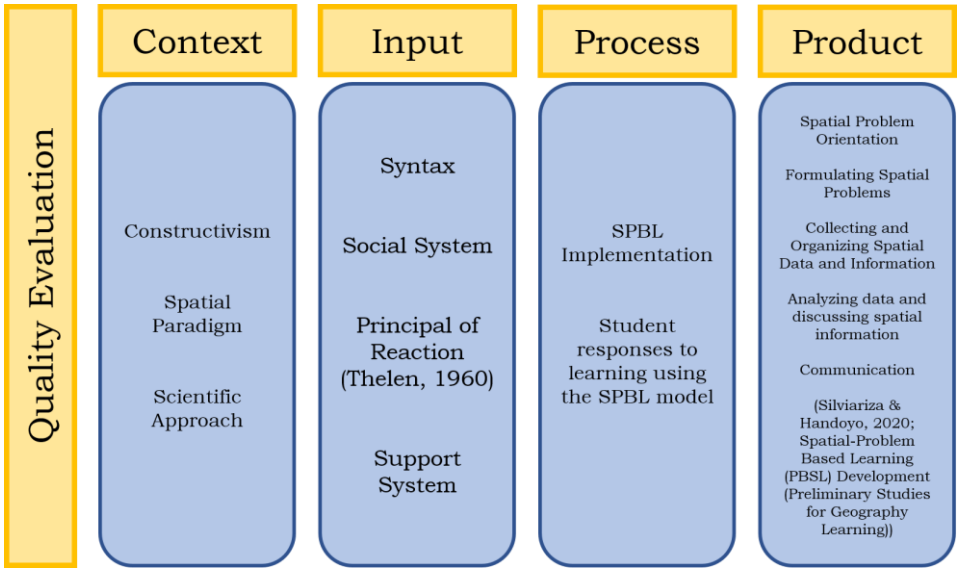

Figure 2
Conceptual framework of the CIPP model for quality evaluation (adaptation from Aziz, Mahmood & Rehman, 2018)

Specifically, context involves identifying a need to decide on the main objectives including the quality of a product/program (Tuna & Başdal, 2021). Input helps determine a responsive project that can handle the identified needs well (Aziz, et al., 2018). Furthermore, the Process monitors the process and potential procedural barriers and identifies deficiencies in the implementation of a program/product (Ellsworth, 2019). Finally, the Product measures, interprets, and evaluates results and interprets the benefits, value, and significance, of a product (Stufflebeam, 2010; Stufflebeam, 2001).

The main strategy is to structure evaluation items while maintaining flexibility. Evaluators/review panelists see design as a process, no longer a product (Zhang, et al, 2011). The goal is to provide a continuous flow of information to the development team to ensure that products developed in a sustainable manner improve their quality.

METHOD

This section describes the sample, the study approach used with the procedures followed to ensure the reliability and validity of the developed product.

Participant

The samples in this study were (1) 22 students of SMA Negeri 1 Pandaan, East Java. The sample is high school students as users of the SPBL 'model of teaching'. Sampling was done randomly. (2) 4 evaluators for product validation of the evaluation instruments that have been made. The evaluator team is the validator who provides advice in terms of curriculum that provides input on construction, content, and language. (3) Expert lecturers to test the readability of the instrument.

Study Approach

The type of research is research and development (R&D). This research and development refer to the 4D research and development step (Thigharajan, et al, 1975). What was developed in this study was the CIPP instrument for the SPBL evaluation model. The stages in this development are: Define, Design, Develop, and Disseminate (Four-D Model).

The stage which is categorized as "Define" is to define a framework to develop an instrument. The next stage "Design" is the design of the instrument prototype and consists of four steps: construction of criteria reference tests, media selection, format selection, and initial design for construction. The "Develop" stage, namely the development process, consists of modifying the prototype material through expert assessment and testing. The final stage of "Disseminate" is summative evaluation, and final packaging activities such as securing copyright and diffusion (Thiagarajan, 1975).

The procedure for developing the CIPP evaluation model instrument in the implementation of the project assessment is following the 4D model adopted with the step of developing a non-test instrument (Rusilowati, 2013).

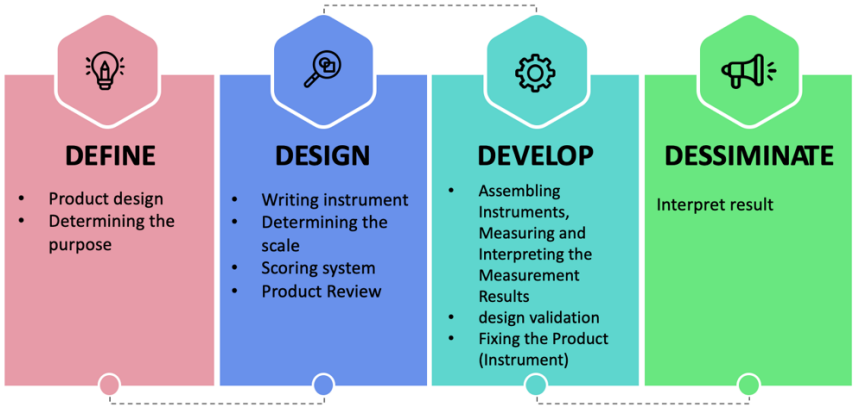

Figure 3
Integration of 4-D Model with Non-Test instrument

Define

The purpose of the non-test instrument developed is to evaluate the SPBL model comprehensively. At this stage, the form of the instrument has been determined, namely an observation sheet, and then determines the selected indicators from the Context, Input, Process, and Product aspects. Items are statements on a scale of 1 to 4.

Design

The writing of the items of the 'model of teaching' evaluation instrument takes into account the theoretical aspects, and the construction and language used in the instrument must be precise and easy to understand.

The instrument scale was used in the development of the instrument. The scoring system carried out by the researcher is the acquisition of scores from observations that have been available for each statement item given by the observer.

Instrument reviews are carried out by expert in the field of development studies. Teachers as evaluators are considered capable of assessing the implementation of a 'model of teaching'. Experts and Evaluators provide assessments and input in the areas of substance, construction, and language in the developed instruments. Students as users provide statements with questionnaires as a form of response to product use.

The trial aims to determine the reliability of the developed instrument. Expert advice and input used will be used as an instrument for improvement.

Develop

The correct statement points are then rearranged into a complete instrument form. Measurements to measure the components of the SPBL model comprehensively. Interpretation of measurements based on CIPP criteria.

Design validation is carried out to determine the extent of the feasibility of an instrument product based on input from experts. Product validation analysis was carried

out descriptively and quantitatively. Quantitative analysis uses Aiken's V analysis (Aiken, 1985) with the following formula:

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

Description:
s = r - lo
n = number of judging panels
lo = lowest validity rating
c = highest validity assessment
r = number given by rater

Furthermore, product reliability analysis used the Intraclass Correlation Coefficient with a 2-way mixed model using the SPSS application. The ICC result is based on a certain index.

Table 1
The reliability index for ICC

ICC Value	Interpretation
0,00 - 0,50	Poor
0,51 - 0,75	Moderate
0,76 - 0,90	Good
0,91 - 1,00	Excellent

Source: Portney and Watkins, 2009

If the ICC value is in the range of 0.00 to 0.50 it means that the reliability of the instrument product is declared low. The ICC value of 0.51 to 0.75 means moderate reliability, if the ICC value is 0.76 to 0.90 then the product reliability can be declared good. The instrument product is declared very well if the ICC value is in the range of 0.91 to 1.00 (Portney and Watkins, 2009).

FINDINGS AND DISCUSSION

The non-test instrument with the CIPP model was developed in the form of a statement (questionnaire) in assessing the SPBL model comprehensively, from model construction to model implementation in learning activities.

Context Evaluation

In this development research, the context is evaluated through theory and document analysis based on the needs of the 'model of teaching'. Context evaluation is defined as an assessment of needs, problems, opportunities, and problems that can be addressed in a particular environment (Stufflebeam, 2000b). Discussions on context evaluation are important issues regarding evaluation in the field of education (Warju, 2016). A strong theory becomes the basis for the development of 'model of teaching's. A 'model of teaching' must be based on a strong theory (Dell'Olio, & Donk, 2007; Mitchell, 2014). The goal is that the 'model of teaching' product can provide the learning environment that students need.

Empirical data from previous studies stated that the SPBL model was constructed by looking more deeply at the needs of the geography class. Orientation and problem-solving in geography learning activities are important so a 'model of teaching' framework with a geographic pattern is needed, namely spatial (Silviariza & Handoyo 2020). In learning activities with the SPBL model, the teacher as a facilitator allows students to construct their knowledge and thoughts to get a meaningful learning experience. According to constructivism theory, humans build knowledge and meaning from their experiences (Bada & Olusegun, 2015).

Table 2
The context in the CIPP evaluation instrument for SPBL model

Aspect	Component	Indicator
Context	The suitability of the SPBL model with Theory	The PBL model is constructivist (Saunders, W. L., 1992) • *Hands-on, Investigative Labs* Point a • *Active Cognitive involvement* Points b, c, d • *Group Work* Point e & f • *Higher-Levels Assessment* Point g
	The SPBL model has been coordinated based on a spatial approach	The SPBL model has been coordinated based on a spatial approach
		The SPBL model is relevant to the scientific approach (Ministry of Education and Culture, 2013)

Source: (Research Data, 2021)

Context evaluation assesses the SPBL model based on needs, goals, assets, and problems in the desired environment. Table 2 shows that the 'Context' aspect provides space to assess the SPBL model according to construction and theory. The purpose of context evaluation is to define, identify and address the needs of the target population, identify problems and assess whether the goals are responsive to the desired needs or not (Stufflebeam, 2001; Stufflebeam, 2010).

The context needs of the SPBL 'model of teaching' on the instrument determine the components of the 'model of teaching' that are adapted to constructivism and spatial theory with a scientific learning approach. Furthermore, indicators in SPBL implementation activities are also adjusted to constructivist rules that must be met, namely hands-on learning activities, investigative labs, active cognitive involvement, group work, and higher-levels assessment (Saunders, W. L., 1992). In addition, in its evaluation SPBL must reflect the scientific learning process of geography (Kemendikbud RI, 2013).

Input Evaluation

Input evaluation includes available and available resources to achieve goals and meet needs (Stufflebeam, 2002). The main purpose of this evaluation is to assess and identify

products to provide information that helps in the use of certain strategies. Therefore, the focus on personnel, resources, procedures, and decisions that define the new objectives is concentrated in this evaluation. Developers/researchers can use input evaluation findings to select, refine, review and revise previously adopted procedural plans (Stufflebeam, 2010). After that, it is necessary to ask how goals can be achieved effectively and efficiently (Stufflebeam, 2000b).

Table 3
Input in the SPBL Model CIPP evaluation instrument

Aspect	Component	Indicator
Input	*Syntax*	The model begins by confronting students with a stimulating problem
		The SPBL model syntax is consistently interrelated
	Social System	SPBL model is democratic
		Decisions developed are from or validated by the group experience
	Principle of Reaction (Thelen, 1960).	The teacher's role in group investigation
		Teacher guides in solving problem or task (What is the nature of the problem? What are the factors involved?)
		Teacher guides group management (What information do we need? How do we organize ourselves to get it?
		A teacher conducts meaning to the individual (How do you feel about this conclusion? What would you do differently after knowing it?)
		The teacher supervises the educational activities to get a meaningful experience
	Support System	The learning process provides flexibility for students to seek information and opinions

Source: (Research Data, 2021)

The input evaluation assesses the competing strategies, work plan, and budget to fulfill the assessment and target the user. Developers can use input evaluation findings to select, refine, review and revise previously adopted procedural plans (Stufflebeam, 2010).

In the rules, there are important things that become elements or inputs for the SPBL model. Among (Joyce & Weil, 2009) namely (a) Syntax is a continuous and orderly stage of model activity, (b) Social System is a situation or atmosphere and norms in a model based on democratic processes and group decisions, and (c) Principles of Reaction are a pattern of activities that describes how teachers see and treat students, including how teachers should respond to them, in this case, the teacher's role in learning activities is as a friendly counselor, consultant, and critic and (d) Support System are all facilities, materials, and tools needed to implement a model in which the environment must be able to respond to the various demands of students. SPBL must provide opportunities for students to seek and collect spatial data and information directly in the field. Students are welcome to investigate and contact resource persons from outside the school environment. This kind of environment provides real experience for students in solving problems and providing factual knowledge.

Process Evaluation

The basic purpose of process evaluation is to provide an overview of all activities in the program (Stuffelbeam, 1971). If it is associated with the evaluation of 'model of teaching's, process evaluation refers to the types of activities carried out in the learning stage. This evaluation provides an overview of the teaching and learning process with the SPBL model.

The 'model of teaching' is a description of the overall approach or teaching plan which includes objectives, steps, learning environment, and system settings. In the process, it is important to pay attention to the impact when implementing the model. According to Joyce and Weil (2009) impact is an instructional learning outcome that is achieved directly by directing students to the expected goals. In addition, the impact of accompaniment is that other learning outcomes are produced by a learning process, as a result of creating a learning atmosphere that is experienced directly by students without receiving direct guidance from the teacher.

Table 4
Process in CIPP evaluation instrument of SPBL model

Aspect	Component	Indicator
Process	SPBL Implementation	At the beginning of the lesson, the teacher informs the students about the learning objectives and achievement criteria
		The teacher informs students about the procedures and assessment
		Students carry out learning using the SPBL model
	Student responses to learning using the SPBL model	▪ Assess student attitudes in implementing learning with the SPBL Model ▪ Enable the improvement of students' geography skills (Bednarz, 1994; Heffron &Downs, 2012)
		Students study in groups
		Students dare to ask the teacher and collages

Source: (Research Data, 2021)

Process evaluation will monitor, document, and assess activities. The developer/ researcher uses the findings of the process evaluation to guide and strengthen the activities in a 'model of teaching' product and to document the activities therein. Process evaluation must provide an opportunity for evaluators to know the process of implementing the 'model of teaching'. In the process, the evaluation of the 'model of teaching' should look at the implementation of the applied model and the student's response to it (Haug, & Ødegaard, 2015; Tanti, et al, 2021). Thus, in designing the instrument product, the component used is how the evaluator sees the implementation of the 'model of teaching' and student responses to the 'model of teaching'.

Product Evaluation

Product evaluation i.e. assessing the impact of the product and assessing the reach of the product to the targeted users and the relevant impact on the environment. Product evaluation activities also ensure that the program reaches its intended beneficiaries (Stufflebeam, 2003). Furthermore, a product has then assessed the extent to which the

product has been successfully adapted and applied to other environments. The product of SPBL is the stages of learning activities. At this stage, there are 5 steps which are then reviewed whether the five stages of the SPBL model can be implemented in the desired environment or not.

Table 5
Product in CIPP Evaluation Instrument of SPBL Model

Aspect	Component	Indicator
Product	SPBL Model	Spatial Problem Orientation
		Formulating Spatial Problems
		Collecting and Organizing Spatial Data and Information
		Analyzing data and discussing spatial information
		Communication

Source: (Research Data, 2021)

Product evaluation aims to examine the impact of a product on the targeted user and the environmental impact. Product evaluation activities also ensure that the program's target benefits are reached (Stufflebeam, 2001; Finney, 2020). Furthermore, a product is then assessed if it has been successfully adapted and applied to other environments.

The scoring system in the instrument used a Likert scale of 1 to 4 based on the result. A selection of observations is available for each item given by the evaluator.

Furthermore, the instrument review activity was managed by the experts. The development of the CIPP model evaluation instrument will be validated by experts. The expert then reviewed and provided feedback on the evaluation instrument items that had been prepared. Also, experts in education and instrument development reviewed and provided advice on the substance, construction, and language of the developed instrument.

Instrument Validity and Reliability

The expert validated that the instrument for assessing the feasibility of the CIPP model in terms of substance, construction, and language was feasible to use. Furthermore, the validity of the evaluators was analyzed using a formula to calculate the content validity coefficient of Aiken (V). Table 6 presents the results of the validity of the 4 evaluators on the items on the instrument.

Table 6
The result of instrument product validation for each item

Item	Evaluator				s1	s2	s3	s4	∑S	n(c-1)	CVI	Description
	I	II	III	IV								
item_1a	4	4	4	4	3	3	3	3	12	12	1,00	HIGH
item_1b	4	4	4	4	3	3	3	3	12	12	1,00	HIGH
item_1c	3	3	4	4	2	2	3	3	10	12	0,83	HIGH
item_1d	3	2	2	2	2	1	1	1	5	12	0,42	MODERATE
item_1e	3	3	3	3	2	2	2	2	8	12	0,67	MODERATE
item_1f	2	3	3	2	1	2	2	1	6	12	0,50	MODERATE
item_1g	3	3	2	3	2	2	1	2	7	12	0,58	MODERATE
item_2	4	3	3	3	3	2	2	2	9	12	0,75	MODERATE
item_3a	2	2	2	3	1	1	1	2	5	12	0,42	MODERATE
item_3b	2	2	2	2	1	1	1	1	4	12	0,33	LOW
item_3c	2	2	2	2	1	1	1	1	4	12	0,33	LOW
item_3d	3	2	3	3	2	1	2	2	7	12	0,58	MODERATE
item_4	2	3	2	2	1	2	1	1	5	12	0,42	MODERATE
item_5	3	3	3	4	2	2	2	3	9	12	0,75	MODERATE
item_6	3	4	3	4	2	3	2	3	10	12	0,83	HIGH
item_7	2	3	3	2	1	2	2	1	6	12	0,50	MODERATE
item_8a	2	2	3	2	1	1	2	1	5	12	0,42	MODERATE
item_8b	3	3	2	3	2	2	1	2	7	12	0,58	MODERATE
item_8c	4	4	3	4	3	3	2	3	11	12	0,92	HIGH
item_9	2	2	2	3	1	1	1	2	5	12	0,42	MODERATE
item_10	3	4	4	4	2	3	3	3	11	12	0,92	HIGH
item_11	2	3	4	2	1	2	3	1	7	12	0,58	MODERATE
item_12	3	2	2	4	2	1	1	3	7	12	0,58	MODERATE
item_13	3	3	3	3	2	2	2	2	8	12	0,67	MODERATE
item_14	2	3	4	2	1	2	3	1	7	12	0,58	MODERATE
item_15	3	2	3	4	2	1	2	3	8	12	0,67	MODERATE
item_16	3	3	4	3	2	2	3	2	9	12	0,75	MODERATE
item_17a	3	3	4	2	2	2	3	1	8	12	0,67	MODERATE
item_17b	3	3	4	4	2	2	3	3	10	12	0,83	HIGH
item_17c	3	3	3	3	2	2	2	2	8	12	0,67	MODERATE
item_17d	2	3	1	2	1	2	0	1	4	12	0,33	LOW
item_17e	3	2	2	3	2	1	1	2	6	12	0,50	MODERATE
item_17f	4	3	3	3	3	2	2	2	9	12	0,75	MODERATE
item_18	3	3	2	2	2	2	1	1	6	12	0,50	MODERATE
item_19	3	2	3	3	2	1	2	2	7	12	0,58	MODERATE
item_20	2	2	2	1	1	1	1	0	3	12	0,25	LOW
item_21	3	4	3	3	2	3	2	2	9	12	0,75	MODERATE
item_22	2	3	2	2	1	2	1	1	5	12	0,42	MODERATE
item_23	3	4	2	4	2	3	1	3	9	12	0,75	MODERATE
item_24	4	4	4	4	3	3	3	3	12	12	1,00	HIGH

Table 6 is shown the validity values that V < 0.4 is categorized as low, 0.4 V < 0.8 is categorized as moderate, V = 0.80 - 1.00 is categorized as valid. Furthermore, the value that V > 0.80 is categorized as high validity (Aiken, 1985; Penfield, & Giacobbi, 2004).

Items 17d and 20 have low values. In both cases, researchers and experts decided to revise some language dictions so that the substance, and instrument are more acceptable

and make sense without changing the substance and purpose of the instrument item. This makes it possible to define items in the instrument in a clearer and more precise way and produce more relevant instrument items (Wiersma, 2001). Thus, the expert states that the instrument product can be used.

After going through the process of improvement and validity from the experts, the evaluators agreed that the instrument could be used. Agreement between raters is very taken into account in making decisions about the validity of a product (Aiken, 1980; Aiken, 1985). So, the instrument is ready to be seen in the attachment.

After the evaluators evaluated the instrument product, then the researcher subjects as many as 22 students to use the instrument product to evaluate the 'model of teaching'. Technically, students learn with the SPBL model. Then, students fill out the instrument in the form of a questionnaire. Tabulation of validation results by subject/user is presented in table 7.

Table 7
Tabulation of instrument product validation results for each item by subject

ITEM	SUBJECT/USER (STUDENTS)																						∑S	n(c-1)	V	Desc.
	sj1	sj2	sj3	sj4	sj5	sj6	sj7	sj8	sj9	sj10	sj11	sj12	sj13	sj14	sj15	sj16	sj17	sj18	sj19	sj20	sj21	sj22				
item_1a	4	4	3	3	3	3	3	4	3	3	3	4	4	2	3	4	4	3	4	3	2	2	49	66	0,74	moderate
item_1b	3	3	3	3	4	4	3	3	3	3	3	4	4	2	3	4	4	3	4	2	2	2	47	66	0,71	moderate
item_1c	3	3	2	2	4	4	4	3	3	3	3	4	4	3	3	4	4	3	4	2	3	3	49	66	0,74	moderate
item_1d	4	4	3	3	3	3	3	4	3	4	3	4	4	3	3	4	4	3	4	3	3	3	53	66	0,80	high
item_1e	3	3	4	4	4	4	4	3	3	4	3	4	4	4	4	4	4	2	4	3	4	4	58	66	0,88	high
item_1f	4	3	4	4	4	4	4	3	3	4	3	4	4	4	3	4	4	2	4	3	4	4	58	66	0,88	high
item_1g	3	3	2	2	3	3	3	3	3	3	3	4	4	2	3	4	4	3	4	2	2	2	43	66	0,65	moderate
item_2	4	4	3	3	3	3	4	4	3	4	3	4	4	3	3	4	4	3	4	3	3	3	54	66	0,82	high
item_3a	4	4	3	3	4	4	4	4	3	3	3	4	4	3	3	4	4	4	4	2	3	3	55	66	0,83	high
item_3b	3	3	3	3	4	4	4	3	3	3	3	4	4	3	3	4	4	3	4	3	3	3	52	66	0,79	moderate
item_3c	4	4	3	3	3	3	3	4	3	3	3	4	4	2	3	4	4	3	4	3	2	2	49	66	0,74	moderate
item_3d	3	3	4	4	4	4	4	3	3	3	3	4	4	3	3	4	4	4	4	3	3	3	55	66	0,83	high
item_4	2	2	3	3	2	2	2	2	3	3	3	4	4	2	3	4	4	2	4	2	2	2	38	66	0,58	moderate
item_5	3	3	4	4	3	3	3	3	3	3	3	4	4	2	3	4	4	3	4	2	2	2	47	66	0,71	moderate
item_6	4	4	4	4	4	4	4	4	3	4	3	4	4	4	3	4	3	3	4	4	4	4	61	66	0,92	high
item_7	3	3	4	4	4	4	4	3	3	4	3	4	4	3	3	4	3	2	4	3	3	3	53	66	0,80	high
item_8a	3	3	3	3	3	3	4	3	3	4	3	3	4	3	3	3	3	3	3	3	3	3	47	66	0,71	moderate
item_8b	3	3	3	3	3	3	3	3	3	4	3	3	4	3	3	3	3	3	3	4	3	3	47	66	0,71	moderate
item_8c	3	3	3	3	3	3	3	3	3	4	3	3	3	3	3	3	3	3	3	4	3	3	46	66	0,70	moderate
item_9	4	4	4	4	4	3	4	4	3	4	3	3	3	3	3	3	3	4	3	3	3	3	53	66	0,80	high
item_10	2	2	3	3	4	4	4	2	3	4	3	3	3	2	3	3	4	3	3	2	2	2	42	66	0,64	moderate
item_11	3	3	3	3	3	3	4	3	3	4	3	3	3	2	3	3	4	3	3	3	2	2	44	66	0,67	moderate
item_12	4	4	3	3	4	4	3	4	3	4	3	4	4	3	3	4	4	3	4	3	3	3	55	66	0,83	high
item_13	3	3	4	4	4	4	3	3	3	4	3	4	4	4	3	4	4	4	4	2	4	4	57	66	0,86	high
item_14	3	3	3	3	4	4	4	3	3	3	3	4	4	3	3	4	4	3	4	3	3	3	52	66	0,79	moderate
item_15	2	2	3	3	2	2	3	2	3	3	3	4	4	2	3	4	4	1	4	2	2	2	38	66	0,58	moderate
item_16	3	3	3	3	3	3	3	3	3	4	3	4	4	3	3	4	4	3	4	3	3	3	50	66	0,76	moderate

The tabulation of the subject/user shows that the value of V is 0.78 (Medium). This value is above the minimum value of V Aiken's which is > 0.4. So, it can be concluded that the questionnaire instrument can be used. The following presents a summary of the validation results by evaluators and users

Table 8
Summary of validation results by evaluators and users of instrument products

item		∑S	V	description
1a-24	Evaluator	300	0,63	moderate
	User	2056	0,78	moderate

Table 8 shows the average value of V. This is a reference that the value of V is above the minimum value of Aiken's V, which is > 0.4, namely 0.63 from evaluators and 0.78 from users. Thus, the evaluation instrument meets the valid criteria. The next criterion that must be met is the reliability of the instrument product. The results of testing the reliability of the instrument product are presented in Table 9.

Table 9
The result of the intraclass correlation coefficient

	Intraclass	95% Confidence Interval		F Test with True Value 0			
	Correlation	Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.472[a]	.312	.635	4.510	39	117	.000
Average Measures	.781[c]	.645	.874	4.510	39	117	.000

Table 8 is shown that the estimated reliability coefficient is in the 'Good' category (0.781). It means that the CIPP evaluation instrument developed is good among evaluators.

FINDINGS

There are obstacles in the process of developing instrument items. It was found that the results of the questionnaire analysis stated items 17d and 20 with a 'low' value. On items with low descriptions, researchers and experts decided to revise some language dictions so that the substance, and instrument are more acceptable and make sense without changing the substance and purpose of the instrument item.

Expert answers during the interview: "…*silahkan untuk merubah diksi, hanya diksi saja yang perlu dirubah tanpa menghilangkan substansinya*".

"…please change the diction, only the diction needs to be changed without losing the substance".

Interview activities were carried out by calling the WhatsApp application and the revision process was carried out with the help of google documents. Both are done simultaneously. Qualitative expert validation is important in developing and improving instrument items. Expert validation aims to improve wording and clarify concepts and substance to avoid uncertainty in item creation (García-Ceberino, et al., 2020). This activity is also made easier because there are not many items that need to be re-agreed after being repaired. This makes it possible to define items in the instrument in a clearer and more precise way and produce more relevant instrument items (Wiersma, 2001). Thus, the expert states that the instrument product can be used.

Table 10
Before and after revision on item description "Low"

No	Statement on items before expert revision	Statement on items after expert revision
17d	Allows students to provide analysis of information and data in solving hypotheses	Allows students to analyze data to answer questions or solve problems
20	Student learning activities in problem orientation can make students formulate spatial problems	Student learning activities in problem orientation can ensure students recognize the problems that occur and formulate them spatially

After going through the revision process and the validity of the experts, the researcher then provided information to the evaluators with the aim that the evaluators gave opinions about the items that had been revised.

Evaluators' answers during the interview: "… *jadi, kami sepakat kalau instrumen yang ini (instrumen setelah revisi ahli) bisa dipakai dilapangan, karna ini bahasanya lebih detail dan jelas*".

"… so, we agree that this instrument (the instrument after expert revision) can be used in the field because this language is clearer and more detailed".

The agreement that the instrument product is suitable for use after the evaluators have assessed the revised instrument item. Agreement between raters is very taken into account in making decisions about the validity of a product (Aiken, 1980; Aiken, 1985). So, the instrument is ready to be seen in the attachment.

CONCLUSION

The purpose of this study was to develop an evaluation instrument. Evaluation is a process that is responsible for monitoring the progress of a product with the desired goals and objectives. In this case, an evaluation instrument will be used to evaluate the SPBL model. To be comprehensive, the SPBL model evaluation instrument was constructed using the CIPP model.

According to different studies, the CIPP model is an effective model used to improve and assess the quality of every corner of the world of education and learning. Many researchers apply the CIPP model to evaluate the quality of textbooks, curricula, and school evaluations. This is because, the CIPP model can evaluate a product as a whole from the point of view of the context, input, process, and product itself.

The procedure for developing the SPBL model evaluation instrument with CIPP follows the 4-D research and development stage. Construction of evaluation instruments with CIPP based on theory and the need for a 'model of teaching'. In its context, the SPBL model was developed with the concept of constructivism with a spatial angle. The input of the SPBL model is properly constructed as a 'model of teaching' in which there are components of syntax, social system, principal of reaction, and support system. The SPBL process can be seen from its implementation activities and student responses

when learning with the SPBL model. Thus, the SPBL product, namely syntax, can be evaluated for its feasibility comprehensively with CIPP.

The output of this study is a good, valid and reliable SPBL model evaluation instrument product. The instrument is in the form of 24 statement items in a questionnaire with a scale of 1-4. Based on the calculation results, the product quality has valid and reliable criteria in terms of construction and substance. Thus, all these aspects meet the criteria, and the product in the form of an evaluation instrument can be used to evaluate the 'model of teaching', especially the SPBL model.

This research provides benefits to the world of education. By looking at the output of this study in the form of an evaluation instrument for 'model of teaching's, teachers can use it to evaluate the 'model of teaching' it uses. This evaluation instrument becomes an evaluation tool for 'model of teaching' developers to provide a direct assessment of the 'model of teaching' they have developed. With this study, it is hoped that evaluation instruments for other 'model of teaching's will emerge. The goal is that the quality of the 'model of teaching' can continue to be corrected and become better.

SUGGESTION

Through this development research, there are several recommendations for education practitioners (teachers) to provide options in implementing a good 'teaching model' to improve the quality of education. In addition, the output of this research in the form of a questionnaire can be used as an evaluation tool for 'model of teaching's with several modifications according to their needs.

Furthermore, this research will be an example for further research on the systematic evaluation of the quality of 'model of teaching's. To be more perfect, it is recommended to analyze the study of concept, substance, language, and content separately and in-depth for each item. In addition, it is recommended to use SEM or CFA analysis with more respondents the better.

ACKNOWLEDGMENT

The authors wish to thank the many persons who contributed their assistance, counsel, and judgment during the writing, field testing, and revision of this paper. Research has no intention or conflict of interest towards individuals or groups.

REFERENCES

Agustina, N. Q., & Mukhtaruddin, F. (2019). The CIPP Model-Based Evaluation on Integrated English Learning (IEL) Program at Language Center. *English Language Teaching Educational Journal*, *2*(1), 22-31.

Aiken, L. R. (1980). Content validity and reliability of single items or questionnaires. *Educational and psychological measurement*, *40*(4), 955-959.

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and psychological measurement*, *45*(1), 131-142.

Al-Shanawani, H. M. (2019). Evaluation of self-learning curriculum for kindergarten using Stufflebeam's CIPP Model. *SAGE Open*, *9*(1), 2158244018822380.;

Arifin, Z. (2016). *Evaluasi Pembelajaran* (Prinsip, Teknik, dan Prosedur) [Learning Evaluation (Principles, Techniques, and Procedures)]. Bandung: Remaja Rosdakarya.

Arikunto, S. & Safrudin, C. (2014). Evaluasi Program Pendidikan Edisi Kedua [Evaluation of Education Program: 2nd Edition]. Jakarta: Bumi Aksara.

Arikunto, S., & Jabar, C. S. A. (2009). Evaluasi Program Pendidikan (2nd ed.) [Evaluation of Education Program: 2nd Edition]. Jakarta: Bumi Aksara.

Asadi, M., Kiany, G. R., Akbari, R., & Samar, R. G. (2016). Program evaluation of the New English Textbook (prospect 1) in the Iranian Ministry of Education. *Theory and Practice in Language Studies*, *6*(2), 291

Asfaroh, J.A., Rosana, D. & Supahar, (2017). Development of CIPP Model of Evaluation Instrument on the Implementation of Project Assessment in Science Learning. *International Journal of Environmental and Science Education, 12*(9).

Aslan, M., & Uygun, N. (2019). Evaluation of Preschool Curriculum by Stufflebeam's Context, Input, Process, and Product (CIPP) Evaluation Model. *Education & Science/Egitim ve Bilim*, *44*(200).

Aziz, S., Mahmood, M., & Rehman, Z. (2018). Implementation of CIPP model for quality evaluation at school level: A case study. *Journal of Education and Educational Development*, *5*(1), 189-206.

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, *5*(6), 66-70

Basaran, M., Dursun, B., Gur Dortok, H. D., & Yilmaz, G. (2021). Evaluation of Preschool Education Program According to CIPP Model. *Pedagogical Research*, 6(2)

Bednarz, S. W. (1994). *Geography for Life: National Geography Standards, 1994*. National Geographic Society, PO Box 1640, Washington, DC 20013-1640.

Bukit, A. V., Bastari, A., & Putra, G. E. (2019). Evaluation of learning programs in Indonesian Naval Technology College with the context, input, process, and product (CIPP) model. *International Journal of Applied Engineering Research*, 14(20), 3823-3827.

Dell'Olio, J. M., & Donk, T. (2007). *Models of teaching: Connecting student learning with standards*. Sage Publications.

Dick, W., Carey, L., & Carey, J. O. (2013). A model for the systematic design of instruction. *Instructional Design: International Perspectives: Theory, Research, and Models*, *1*, 361-370.

Ebtesam, E., & Foster, S. (2019). Implementation of CIPP Model for Quality Evaluation at Zawia University. *International Journal of Applied Linguistics and English Literature*, *8*(5);

Ellsworth, R. M. (2019). *Using Curriculum Mapping to Identify Improvements in an On-The-Job Training Program: Initiating a Program Evaluation Using the CIPP Evaluation Model* (Doctoral dissertation, University of Florida).

Finney, T. L. (2020). Confirmative Evaluation: New CIPP Evaluation Model. *Journal of Modern Applied Statistical Methods*, *18*(2), 30.

Galvin, J. C. (1983). What can trainers learn from educators about evaluating management training? *Training and Development Journal, 37*(8), 52–57.

García-Ceberino, J. M., Antúnez, A., Ibáñez, S. J., & Feu, S. (2020). Design and Validation of the Instrument for the Measurement of Learning and Performance in Football. *International Journal of Environmental Research and Public Health*, *17*(13), 4629.

Hakan, K., & Seval, F. (2011). CIPP evaluation model scale: development, reliability, and validity. *Procedia-Social and Behavioral Sciences*, *15*, 592-599.

Haug, B. S., & Ødegaard, M. (2015). Formative assessment and teachers' sensitivity to student responses. *International Journal of Science Education*, *37*(4), 629-654.

Heffron, S. G., & Downs, R. M. (2012). *Geography for Life: National Geography Standards; Geography Education Standards Project (GENIP)*. National Council for Geographic Education.

Iqbal, Z., Anees, M., Khan, R., Wadood, A., & Malik, S. (2021). A Comparative Analysis of the Efficacy of Three Program-Evaluation Models–A Review on their Implication in Educational Programs. *Humanities & Social Sciences Reviews*, *9*(3), 326-336.

Joyce, B., & Weil, M., & Calhoun, E. (2003). Models of Teaching (7th ed.). Boston: Allyn & Bacon.

Joyce, B., Weil, M., & Calhoun, E. (2009). Models of teaching (8th ed.). Boston, MA: Pearson Education.

Kementerian Pendidikan dan Kebudayaan (Ministry of Education and Culture). 2013. Permendikbud Nomor 65 Tahun 2013 Tentang Standar Proses Pendidikan Dasar dan Menengah [Regulation of the Ministry of Education and Culture About Standards for Primary and Secondary Education Processes]. Jakarta: Kementerian Pendidikan dan Kebudayaan RI.

Kurnia, F., Rosana, D., & Supahar. (2017, August). Developing evaluation instruments based on CIPP models on the implementation of portfolio assessment. *In AIP Conference Proceedings* (Vol. 1868, No. 1, p. 080003). AIP Publishing LLC.)

Lippe, M., & Carter, P. (2018). Using the CIPP model to assess nursing education program quality and merit. *Teaching and Learning in Nursing*, *13*(1), 9-13

Mitchell, D. (2014). *What really works in special and inclusive education: Using evidence-based teaching strategies*. Routledge

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. *Design Research*, *153*, 152-169.

Penfield, R. D., & Giacobbi, Jr, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in physical education and exercise science*, *8*(4), 213-225.

Peter, Olivia, F. (1992). *Developing Curriculum*. New York: Harper Collins Publisher.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice* (Vol. 892). Upper Saddle River, NJ: Pearson/Prentice Hall.

Rachmaniar, R., Yahya, M., & Lamada, M. (2021). Evaluation of Learning through Work Practices Industry Program at University with the CIPP Model Approach. *International Journal of Environment, Engineering, and Education*, *3*(2), 59-68.

Rusilowati, A. (2013). Pengembangan Instrumen Nontes (Non-test Instrument Development). In *Makalah. Seminar Nasional Evaluasi Pendidikan Di Universitas Negeri Semarang* (pp. 7-21).

Sanjaya, M. E. (2015). Pengembangan Instrumen Evaluasi pada praktikum Uji Enzim Katalase di SMA Negeri Titian Teras Muaro Jambi [Development of Evaluation Instruments in the Catalase Enzyme Test practicum at SMA Negeri Titian Teras Muaro Jambi]. *Edu-Sains* 4, 2.

Sanusi, A., Maulana, D., & Sabarno, R. (2021, December). Evaluation of Students Writing Skills Using CIPP model in Arabic Learning: The Concept and its Implementation. *In International Conference on Arabic Language and Literature* (pp. 207-214).)

Saunders, W. L. (1992). The constructivist perspective: Implications and teaching strategies for science. *School science and mathematics*, 92(3), 136-141.

Silviariza, W. Y., & Handoyo, B. (2020). Spatial problem-based learning (SPBL) development (preliminary studies for geography learning). *Jurnal Pendidikan Geografi: Kajian, Teori, dan Praktik dalam Bidang Pendidikan dan Ilmu Geografi*, *25*(1), 69-79.

Silviariza, W. Y. Sumarmi., & Handoyo, B.(2021). Improving Critical Thinking Skills of Geography Students with Spatial Problem Based Learning (SPBL). *International Journal of Instruction*, *14*(3), 133-152.

Stufflebeam, D. (1985). *Systematic Evaluation "A Self-Instruction Guide to Theory and Practice*". Hingham Massachusetts: Kluwer Academic Publishers.

Stufflebeam, D. (2001). Evaluation models. *New directions for evaluation, 2001*(89), 7-98.

Stufflebeam, D. L. (2000a). The CIPP model for evaluation. In *Evaluation models* (pp. 279-317). Springer, Dordrecht.

Stufflebeam, D. L. (2000b). Foundational models for 21 st century program evaluation. In *Evaluation models* (pp. 33-83). Springer, Dordrecht.

Stufflebeam, D. L. (2002). Institutionalizing evaluation checklist. *The Evaluation Center: Western Michigan University. Retrieved March 18*, 2007.

Stufflebeam, D. L. (2003). The CIPP model for evaluation. In D. L. Stufflebeam & T. Kellaghan (Eds.), *The international handbook of educational evaluation* (Chapter 2). Boston, MA: Kluwer Academic Publishers.

Stufflebeam, D. L. (2010). CIPP evaluation model checklist: A tool for applying the CIPP model to assess projects and programs. San Francisco, CA: Jossey Bass.

Stufflebeam, D. L., & Coryn, C. L. (2014). *Evaluation theory, models, and applications* (Vol. 50). John Wiley & Sons.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models, & applications. San Francisco, CA: Jossey-Bass.

Stufflebeam, D. L., & Shinkfield, A. J. (2012). *Systematic evaluation: A self-instructional guide to theory and practice* (Vol. 8). Springer Science & Business Media.

Tanti, T., Kurniawan, D. A., Sukarni, W., Erika, E., & Hoyi, R. (2021). Description of Student Responses Toward the Implementation of Problem-Based Learning in Physics Learning. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, *6*(1), 30-38.

Thelen, H. A. (1960). Education and the Human Quest: What's to Become of Johnny? *The School Review*, *68*(2), 136-151.

Thiagarajan, Sivasailam, Semmel, Dorothy S., Semmel, Melvyn I. (1975). *Instructional Development for Training Teachers of Exceptional Children*. Bloomington, Indiana: Indiana University.

Tiantong, M., & Tongchin, P. (2013). A multiple intelligences supported web-based collaborative learning model using Stufflebeam's CIPP evaluation model. *International Journal of Humanities and Social Science*, *3*(7), 157-165.

Tuna, H., & Başdal, M. (2021). Curriculum evaluation of tourism undergraduate programs in Turkey: A CIPP model-based framework. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 29, 100324.

Umam, K. A., & Saripah, I. (2018). Using the Context, Input, Process, and Product (CIPP) model in the evaluation of training programs. *International Journal of Pedagogy and Teacher Education*, 2, 19-183.

Warju, W. (2016). Educational program evaluation using CIPP model. *INVOTEC*, *12*(1)

Wiersma, L. D. (2001). Conceptualization and development of the sources of enjoyment in youth sport questionnaire. *Measurement in Physical Education and Exercise Science*, *5*(3), 153-177.

Yusuf, A. M. (2017). *Asesmen dan evaluasi pendidikan* [Educational assessment and evaluation]. Prenada Media.

Zhang, G., Zeller, N., Griffith, R., Metcalf, D., Williams, J., Shea, C., & Misulis, K. (2011). Using the context, input, process, and product evaluation model (CIPP) as a comprehensive framework to guide the planning, implementation, and assessment of service-learning programs. *Journal of Higher Education Outreach and Engagement*, *15*(4), 57-84.

ATTACHMENT

CIPP Evaluation Instrument of SPBL

Aspect	Component	Indicator	No	Description	1	2	3	4
Context	The suitability of the SPBL model with Theory Allows students to think spatially a. Seeing phenomena with facts, b. prejudice-free, c. the objective on a spatial problem, d. analyzing spatial data and information	The PBL model in constructivism (Saunders, W. L., 1992) ***Hands-on, Investigative Labs*** Point a ***Active Cognitive involvement*** Points b, c, d ***Group Work*** Point e, f ***Higher-Levels Assessment*** Point g	1	Allows students to learn and find immediate answers to the questions made				
				Allows students to interpret data				
				Allows students to participate in cognitive conflict (constructive argumentation about the phenomenon under study)				
				Allows students to develop and select plausible hypotheses to complete explanations				
				Allows students to work with teams				
				Allows students to discuss with the team				
				Allows students to think at a higher cognitive level				
	The SPBL model has been coordinated based on a spatial approach	The SPBL model has been coordinated based on a spatial approach	2	Allows students to think spatially				
		The SPBL model is relevant to the scientific approach (Ministry of Education and Culture, 2013)	3	Seeing phenomena with facts,				
				prejudice-free,				
				objective on a spatial problem,				
				analyzing spatial data and information				
Input	*Syntax*	The model begins by confronting students with a stimulating problem	4	The model begins by confronting students with a stimulating problem				
		The SPBL model syntax is consistently interrelated	5	The SPBL model syntax is consistently interrelated				
	Social System	SPBL model is democratic	6	SPBL model is democratic				
		Decisions developed are from or validated by the group experience	7	Decisions based on experience and group deliberation				

		The teacher's role in group investigation	8	Allows teachers to act as counselors
				Allows teachers to act as consultants
				Allows the teacher to act as a friendly critic
		Teacher guides in solving problem or task (What is the nature of the problem? What are the factors involved?)	9	Allows teachers to guide problem-solving
	Principle of Reaction	Teacher guides group management (What information do we need? How do we organize ourselves to get it?	10	Allows teachers to guide group management
		The teacher conducts meaning to the individual (How do you feel about this conclusion? What would you do differently after knowing it?)	11	Enables teachers to convey meaning to individuals
		The teacher supervises the educational activities to get a meaningful experience	12	Allows the teacher to supervise the educational activities to get meaningful experience
	Support System	The learning process provides flexibility for students to seek information and opinions	13	Allows students to freely seek information and opinions outside the classroom
Process	*SPBL Implementation*	At the beginning of the lesson, the teacher informs the students about the learning objectives and achievement criteria	14	Allows the teacher to inform students about the learning objectives and achievement criteria
		The teacher informs students about the procedures and assessment	15	Allows the teacher to inform students about the procedures and types of assessment to be used
		Students carry out learning using the SPBL model	16	Allows students to carry out learning according to the SPBL model syntax
	Student responses to learning using the SPBL model	Assess student attitudes in implementing learning with the SPBL Model Enable the improvement of students' geography skills (Bednarz, 1994; Heffron & Downs, 2012):	17	Allows students to identify problems and ask geographic questions
				Allows students to collect data including observations and measurements of geographic phenomena
				Allows students to organize or process data
				Allows students to analyze data to answer questions or solve problems
				Allows students to answer or solve problems
				Allows students to communicate or inform geographic data to audiences such as teachers and collages
		Students study in groups	18	Allows students to learn and work with teams

		Students dare to ask the teacher and collages	19	Enables communication between teachers and students
Product	SPBL Model	Spatial Problem Orientation (definition of problems, roles, tasks, and others)	20	Student learning activities in problem orientation can ensure students recognize the problems that occur and formulate them spatially
		Formulating Spatial Problems	21	The teacher ensures that student learning activities are carried out in an organized, both individually and in groups
		Collecting and Organizing Spatial Data and Information (Individual and Group Study)	22	Student learning activities in data collection and organization can ensure data will be collected and spatially organized
		Analyzing data and discussing spatial information	23	Student activities in analyzing data and discussing answers can ensure data is analyzed and answered spatially
		Communication	24	Student activities in communicating can ensure the results will be communicated effectively