Validation of Reflective Model Curriculum Evaluation Instruments

Istyarini Istyarini, Niam Wahzudik, Wardi Wardi, Nurussa'adah Nurussa'adah and Christian Christian
Department of Curriculum and Educational Technology
Faculty of Education, Universitas Negeri Semarang, Indonesia
istyarinitp@mail.unnes.ac.id, niam.wahzudik@mail.unnes.ac.id,
wardifip@mail.unnes.ac.id, nurussaadah@mail.unnes.ac.id
tian@mail.unnes.ac.id

Abstract

The success of curriculum evaluation activities is determined by the evaluation tools or instruments used in collecting evaluation data. The absence of quality basic education curriculum evaluation instruments urgently needs to develop a good and quality curriculum evaluation instrument. A good evaluation instrument must meet the requirements of the instrument being valid and reliable. This research aimed is producing valid and reliable curriculum evaluation instruments and guides. The research method used in this research is development research. The data was collected using a questionnaire technique, namely the expert validation sheet. The analysis was carried out quantitatively. The results of the second phase of research in 2020 show that the results of expert validation show that from the aspect of content feasibility, the percentage of feasibility is 97.62% with very feasible criteria. The linguistic feasibility aspect obtained a percentage of 83.33% with Eligible criteria. Whereas in the presentation aspect, it obtained a feasibility percentage of 94.44% which means it is very feasible. This means that the curriculum evaluation instrument of this development is quality with this curriculum evaluation instrument, it is hoped that it can make it easier for education units, especially elementary school, to carry out a comprehensive curriculum evaluation.

Keywords:
Instrument development, Curriculum evaluation, Reflective model, Education unit curriculum, Elementary school curriculum

1. Introduction

Curriculum evaluation activities play an important role both in determining educational policy in general, and in making decisions in the curriculum. The results of curriculum evaluation can be used by education policy holders and curriculum developers in selecting and establishing policies for developing the education system and developing the curriculum model used. The results of curriculum evaluation can also be used by teachers, school principals and other education implementers, in understanding and fostering student development, selecting learning materials, selecting learning methods and tools, assessment methods and other educational facilities.

Legal formal curriculum evaluation activities have been regulated in Law Number 20 of 2003 concerning the National Education System articles 55 and 56 which are then further regulated by the issuance of Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 159 of 2014 concerning Curriculum Evaluation. The issuance of the Permendikbud provides an opportunity to carry out curriculum evaluation activities in each school. In general, when discussing evaluation activities in schools it easily leads to learning outcomes assessment activities as a form of curriculum results, whereas ideally curriculum evaluation activities can be carried out on curriculum ideas and designs, curriculum documents, curriculum (implementation) processes, and results (impact). curriculum. The phenomenon of limited knowledge, confusion of understanding and the background of the competence of school human resources has resulted in a narrowing of the concept and practice of curriculum evaluation in schools which is actually very broad.

The reality in the field also shows that schools do not yet have specific instruments that are used to evaluate the curriculum from various curriculum dimensions including curriculum ideas, curriculum documents, curriculum
implementation and the results and impacts of the curriculum applied by the school. A good evaluation activity of course must use evaluation instruments that are also good and of quality, such as validity and reliability, practicality, and effectiveness to measure what will be evaluated.

In accordance with the results of Phase 1 research (2019), it shows that curriculum evaluation activities at the basic education level have generally been carried out in each school even though in terms of quality between one school and another school, the process achievement level and evaluation results are quite diverse. This is because in every school there is no standard and specific curriculum evaluation instrument that can be used by schools. Usually the curriculum evaluation activities using school self-evaluation (EDS) according to the schools that use it include instruments that are not easy to use because some items from the EDS instrument are not easy for schools to understand. To support curriculum evaluation activities in elementary schools, schools need curriculum evaluation instruments that are easy to implement in the field and need to hold training activities for curriculum preparation and evaluation with a minimum implementation of 32 hours.

Many previous research activities are related to curriculum evaluation, for example research conducted by (Dewanti & Widyasari, 2017) whose research results show that the 2013 curriculum learning evaluation instrument based on multimedia on the Physics material of Class X SMA is in good category (83.7%) so that the learning evaluation instrument Curriculum 2013 based on digital media is not revised, it can be implemented in schools. Research by (Budiani et al., 2017) on the evaluation of the implementation of the 2013 curriculum at Independent Implementing Schools shows that the 2013 curriculum can run very well with the support of meeting national education standards and teachers who have good motivation, creativity, and performance.

(Setiadi, 2016) research raised the theme of the implementation of assessment in the 2013 Curriculum, showing the results that several things need to be improved in the implementation of learning outcomes assessment in the 2013 curriculum, which includes improvements in socialization and training in making assessment tools, simplifying guidelines for assessing K13 learning outcomes, and using range assessments 1-4 on the knowledge and skills assessment. Research that is almost similar is the research conducted by (Gultom et al., 2014) which examines the development of a reflective evaluation model of the English education curriculum in the development of National Character which shows that the evaluation model of development results is valid and reliable in the form of evaluation guides.

The four previous studies above show that curriculum evaluation activities have been carried out in schools, but the curriculum evaluation activities are more likely to be carried out on the object of evaluation (evaluation), namely the dimensions of curriculum implementation shown by studies conducted by Budiani et al, and Setiadi. Meanwhile, studies related to the development of evaluation instruments carried out by Dewanti and Widyasari were only applied to one subject, namely SMA class X physics. The fourth research by Gultom et al. Has produced a reflective evaluation model of the English education curriculum and its application in higher education (PT).

Referring to previous studies, the research that will be carried out in this research is to attempt to produce reflective evaluation instruments and guides that can be used by schools, especially elementary schools (SD) to evaluate the curriculum they have developed, applied and they evaluate its success which is not only in aspects learning outcomes but from the dimensions of ideas, documents (KTSP documents), implementation processes, and products (results).

Based on these conditions, it is necessary to develop curriculum evaluation instruments along with a curriculum evaluation guidebook that can be used as a source of reference in any curriculum evaluation activity because what is stated in the Regulation of the Minister of Education and Culture of the Republic of Indonesia Number 159 of 2014 concerning Curriculum Evaluation has not been technically regulated. evaluation of the curriculum in the field. With the existence of valid and reliable evaluation instruments, it is expected that each curriculum evaluation activity will be able to produce objective information on the curriculum implemented in schools so that it can produce decision recommendations for curriculum policy makers in each school. The purpose of this study is to determine the validity (feasibility) of the curriculum development evaluation instrument.
1.1 Objectives
Based on the background of the problem above, the research objectives was to validate and conduct limited trials of the developed curriculum evaluation instruments.

2. Literature Review
There are various opinions from curriculum leaders in defining curriculum evaluation. The following are some definitions of curriculum evaluation from curriculum leaders / experts. (Hasan, 2008) states that curriculum evaluation is a systematic effort to collect information about a curriculum to be used as a consideration of the value and meaning of the curriculum in a particular context. According to (Gronlund, 1985) curriculum evaluation formulation is a systematic process of collecting, analyzing and interpreting information / data to determine the extent to which students have achieved learning objectives. According to (Tyler, 1949) evaluation focuses on efforts to determine the level of changes that occur in learning outcomes. Meanwhile, Morrison believes that evaluation is an act of judgment based on a set of agreed and accountable criteria.

The purpose of curriculum evaluation varies depending on the concept or understanding held by the evaluator and others about the evaluation concept. The objectives of curriculum evaluation are divided into the following: (1). Providing information regarding the implementation of the development and implementation of a curriculum as input for decision making, (2). Determine the level of success and failure of a curriculum and the factors that contribute to a certain environment, (3). Develop a variety of problem-solving alternatives that can be used in efforts to improve the curriculum, and (4). Understand and explain the characteristics of a curriculum and the implementation of a curriculum. The purpose of curriculum evaluation according to (Tyler, 1949) is to determine the level of change in behavior that occurs, both statistically and educatively.

The evaluation function is categorized into two, namely formative and summative functions. The formative function is an evaluation function to provide information and considerations with regard to efforts to improve a curriculum. This function can only be carried out when the curriculum is still in the development / construction process, curriculum document development has not been completed or is still in the process of being prepared. The formative function of curriculum evaluation can only be performed when the evaluation is concerned with the process rather than focusing on the results. While the summative function is a curriculum evaluation function to give consideration to the results of curriculum development. The results of these developments can be in the form of curriculum documents, learning outcomes, or the impact of the curriculum on schools and communities. From this summative function, the evaluator can give consideration to whether a curriculum needs to be continued, still relevant to the development and demands of society, or vice versa, it must be replaced because of failure and is not relevant to the development and demands of society.

In terms of refining and or improving the curriculum, curriculum evaluation has a function that provides information on why the curriculum should change and be changed. Therefore, it is very open to conduct curriculum evaluations concerning ideas and conceptions which is called reflective evaluation. Reflective evaluation refers to evaluating the dimensions of the curriculum as ideas or conceptions. Curriculum ideas are the most important component in the curriculum development process. The curriculum idea is a formulation of the educational philosophical position adopted, the theoretical view of the curriculum concept, the curriculum model used, the concept of content, curriculum organization, curriculum design, curriculum document design, the position of students in learning. The idea of a curriculum must be clear because it becomes the basis and foundation for the development of various components of the curriculum document (Hasan, 2008).

Based on this, problems will arise if the curriculum document development team and internal evaluators do not understand the curriculum idea. If the curriculum development team does not understand the curriculum idea that is used as the basis for curriculum development it can result in problems related to the relevance of curriculum ideas to community needs, clarity of curriculum ideas, and understanding of ideas by the curriculum development team as well as at the level of implementation and the results and impacts of enforce a curriculum. Effectiveness is a measure that states the extent to which goals or objectives (quantity, quality and time) have been achieved (Sutomo, 2012). Referring to this definition when linked to an assessment instrument, effectiveness is related to quality (Djibu et al., 2019). Quality is a comprehensive description and characteristics of the assessment instrument developed in research with reference to the requirements of a good assessment instrument. A good assessment instrument is an
instrument that meets certain requirements or rules, can provide accurate data in accordance with its function, and only measures certain samples of behavior.

The characteristics of a good assessment instrument are valid, reliable, relevant, representative, practical, discriminative, specific and proportional (Arifin, 2009). Valid, meaning that an instrument can be said to be valid if it is able to measure what it wants to measure accurately. Reliable, meaning that an instrument can be said to be reliable or reliable if the instrument has consistent results. Relevant, meaning that the assessment instrument used must be in accordance with the competency standards, basic competencies, and predetermined indicators (Shofwan et al., 2019). Assessment instruments must be in accordance with the domain of learning outcomes, such as cognitive, affective, and psychomotor domains. Representative, meaning that the assessment instrument material must truly represent all the material presented. For this reason, in developing an assessment instrument, you must use the syllabus as a reference for selecting the material for the assessment instrument.

Practical, meaning that the assessment instrument is easy to use. Practicality is not only seen from the technique of arranging the instrument, but also for other people who want to use the assessment instrument (Siswati et al., 2020). Discriminative, meaning that the instrument must be arranged in such a way that it can show the slightest differences. Specific, meaning that an assessment instrument is prepared and used specifically for the object being assessed. Proportional, meaning that an assessment instrument must have a level of difficulty that is proportional to difficult, medium, and easy. Meanwhile (Kanto et al., 2020; Nuraini et al., 2019; Sugiyanto et al., 2015) say that an evaluation model is said to be effective or good if it meets the requirements: comprehensive, practical, economical, and is supported by valid and reliable data collection instruments, and has clarity of language. The assessment instrument must be accountable.

3. Methods

The research method used in this research is development research, begins with the validation of instrument prototypes and guidance by experts and practitioners to test product feasibility theoretically and product feasibility to be implemented empirically. After going through the process of validation and revision, an initial (limited) field trial of the application of prototype instruments and a reflective evaluation guide for basic education curriculum in schools was carried out in 2 (two) education units representing the level of Elementary School (SD). Data collection was carried out by using a questionnaire technique. The analysis was carried out quantitatively for the data from the questionnaire.

4. Results and Discussion

(Sugiyono, 2010) states, "a valid instrument means the measuring instrument used to obtain data (measure) is valid". Valid means that the instrument measures what it should measure. Validity is used in measuring the instrument to determine the validity level of the instrument being tested. There are several ways to test the validity of the instruments used in the study, namely Construct Validity Testing, Content Validity Testing and External Validity Testing.

The test used in this research is to use testing the validity of construction (Construct Validity). Broadly speaking, according to (Sugiyono, 2010) construct validity can be interpreted as the concept of validity that is constructed regarding the aspects to be measured based on certain theories, which are then consulted with experts. Experts are asked for their opinion regarding the research instruments that have been made, then these instruments can be used without improvement, there are improvements and can be completely overhauled. The results of the validation of the results of the development evaluation instrument are as follows.

<table>
<thead>
<tr>
<th>Aspects assessed</th>
<th>Expert Value Score 1</th>
<th>Expert Value Score 2</th>
<th>Maximum value</th>
<th>Percentage</th>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feasibility of content</td>
<td>20</td>
<td>21</td>
<td>42</td>
<td>97.62%</td>
<td>Very feasible</td>
</tr>
<tr>
<td>Language</td>
<td>8</td>
<td>7</td>
<td>18</td>
<td>83.33%</td>
<td>Eligible</td>
</tr>
<tr>
<td>Presentation</td>
<td>9</td>
<td>8</td>
<td>18</td>
<td>94.44%</td>
<td>Very feasible</td>
</tr>
</tbody>
</table>

Based on the results of expert validation, it shows that from the aspect of content feasibility it shows the feasibility percentage of 97.62% with very feasible criteria. The linguistic feasibility aspect obtained a percentage of 83.33%
with Eligible criteria. Whereas in the presentation aspect, the feasibility percentage is 94.44% which means it is very feasible.

The limited trial phase of this research is intended to find out how the comments, views, opinions and assessments of elementary school teachers on the evaluation instrument of development results. Limited trials were carried out by distributing questionnaires to elementary school teachers. The test results can be presented as follows.

First, in terms of the suitability aspect of the instrument for the purpose of curriculum evaluation, the respondent's assessment provides an assessment that the instrument is in accordance with the objectives of curriculum evaluation. In detail, 90% of respondents said it was good and 10% of other respondents said it was quite good.

Second, on the aspect of the suitability of the instrument rubric with the indicators of the curriculum evaluation process, the views of the respondents indicated that 10% said it was quite appropriate and the majority or 90% of respondents said the instrument rubric was in accordance with the indicators of the curriculum evaluation process.

Third, the respondent's assessment of the aspect of the instrument's ability to develop results in revealing the objective of the evaluation showed that 93% of respondents said it was good, and 7% of other respondents said it was quite good.

Fourth, the respondent's assessment of the ability of the developed instrument in evaluating curriculum ideas proved that 7% of respondents gave a fairly good assessment, while 93% of other respondents stated that the instrument had a good ability to evaluate curriculum ideas. The data can be seen in the following figure.

Fifth, the next aspect assessed by respondents is the aspect of the instrument's ability to evaluate curriculum documents. 3% of respondents stated that their abilities were not good. Other respondents stated that they have a fairly good ability (3%) and 94% of other respondents stated that they have good abilities in evaluating curriculum documents, the research data can be presented in the following figure.

Sixth, the next aspect assessed by the respondent is how the ability of the development results instrument in evaluating the implementation of the curriculum, the results of this study indicate that 7% of respondents said it was good enough and most of the respondents, namely 93% of respondents said it was good and there were no respondents who gave a bad assessment (0%).

Seventh, the aspect of the instrument's ability to evaluate the results of the curriculum was also revealed in this study, the result was that the majority of respondents (93%) stated that the instrument of development results was able to evaluate the results of the curriculum well, while 7% stated that they had quite good abilities.

Eighth, the quality of the writing of the instruments developed by their suitability to the language rules of the research results revealed that 3% of respondents stated that the writing of the instruments was not good, 93% of
respondents gave good ratings and the other 3% said they were quite good. The research data is summarized in the following figure.

![Figure 3. The Writing of the Instruments Based on the Language Rules](image)

Tenth, relating to the ease with which respondents understand the language used in the development results instrument shows that 90% of respondents stated that the language of the instrument was well understood, 7% said it was good enough, and 3% said it was not good or difficult to understand.

![Figure 4. The Ease of Instrument's Language to be Understood](image)

Eleventh, the next aspect assessed by the respondent is whether the instrument of development results is able to clearly present the evaluation objectives. The results of this study prove that 87% of the instruments serve the evaluation objectives well, 10% are good enough and the other 3% state that the objectives are not good or clear. evaluations listed on the instrument of development results.

Evaluation is an act in consideration based on a set of agreed and accountable criteria (Morrison), of which there are 3 main factors, namely judgment, then a description of the object of assessment, and the criteria that can be accounted for. (Hasan, 2008) states that curriculum evaluation is a systematic effort to collect information about a curriculum to be used as a consideration of the value and meaning of the curriculum in a particular context. According to (Gronlund, 1985), curriculum evaluation formulation is a systematic process of collecting, analyzing and interpreting information / data to determine the extent to which students have achieved learning objectives. According to (Tyler, 1949) evaluation focuses on efforts to determine the level of change that occurs in learning outcomes. Meanwhile, Morrison believes that evaluation is an act of judgment based on a set of agreed and accountable criteria.

To produce objective assessment data requires a good and quality assessment instrument. One of the requirements for a quality assessment instrument is that the assessment instrument must be valid because a valid instrument will produce valid data as well. An assessment instrument is said to be valid if the instrument can accurately measure what it wants to measure. The term valid is often referred to as valid or appropriate. This means that it is appropriate to measure what is being measured. The standard in (Mardapi, 2012) states that validity or validity refers to the degree of evidence and theory supports the interpretation of test scores as the purpose of using the test. The validation process is the collection of evidence to provide a scientific basis for the interpretation of test scores. Validity, which is evidence and theoretical support for the interpretation of the test in accordance with the objectives of the test, is the most basic foundation in developing and evaluating a test (Mardapi, 2012)

A good assessment instrument is an instrument that meets certain conditions or rules, can provide accurate data according to its function, and only measures certain samples of behavior. The characteristics of a good assessment instrument are valid, reliable, relevant, representative, practical, discriminative, specific and proportional (Arifin, 2009). Valid, meaning that an instrument can be said to be valid if it is able to measure what it wants to measure accurately. Reliable, meaning that an instrument can be said to be reliable or reliable if the instrument has consistent results. Relevant, meaning that the assessment instruments used must be in accordance with the competency standards, basic competencies, and predetermined indicators. Assessment instruments must be in accordance with
the domain of learning outcomes, such as cognitive, affective, and psychomotor domains. Representative, meaning that the assessment instrument material must truly represent all the material presented. For this reason, in developing an assessment instrument, you must use the syllabus as a reference for selecting the material for the assessment instrument.

Practical, meaning that the assessment instrument is easy to use. Practicality is not only seen from the instrument preparation technique, but also for other people who want to use the assessment instrument. Discriminative, meaning that the instrument must be arranged in such a way that it can show the slightest differences. Specific, meaning that an assessment instrument must have a level of difficulty that is proportional to difficult, medium, and easy. Meanwhile (Sugiyanto et al., 2015) say that an evaluation model is said to be effective or good if it meets the requirements: comprehensive, practical, economical, and is supported by valid and reliable data collection instruments, and has clarity of language. The assessment instrument must be accountable.

6. Conclusion
The results of expert validation show that from the aspect of content feasibility, the percentage of feasibility is 97.62% with very feasible criteria. The linguistic feasibility aspect obtained a percentage of 83.33% with Eligible criteria. Whereas in the presentation aspect, it obtained a feasibility percentage of 94.44% which means it is very feasible. The limited trial results showed that the majority or 92% of respondents stated that they had good quality, 7% were good enough and only a small part, namely 1% of respondents, stated that they were not good. This means that the curriculum evaluation instrument of this development is quality.

References
**Biography**

**Istyarini** is associate professor in the curriculum and tenology department at the Faculty of Education at Universitas Negeri Semarang. Obtained a bachelor's degree in Education Curriculum and Technology, Yogyakarta State University, graduated from a master's degree in education management at Universitas Negeri Semarang. She is interested in researching the field of educational evaluation and learning media. Many of his writings were published in journals and seminar proceedings. Currently she is a member of the Association of Indonesian Educational Technology Study Programs (APSTPI).

**Niam Wahzudik** is an assistant professor in the Department of Curriculum and Educational Technology, Universitas Negeri Semarang. Graduated from a Bachelor of Education in the Educational Technology study program at Universitas Negeri Semarang and a Masters in Curriculum Development at Universitas Negeri Semarang. He has published his writings in journals and seminar proceedings. His research interests are in the field of curriculum development and evaluation studies, blended learning, learning models, and innovation in the field of educational technology. He is a member of the Association of Indonesian Educational Technology Study Programs (APSTPI) and the Indonesian Curriculum Developers Association (HIPKIN).

**Wardi** is assistant professor in the Department of Curriculum and Educational Technology, Universitas Negeri Semarang. Graduated from a Bachelor of Education in the Educational Technology study program at Universitas Negeri Semarang and a master’s in education management at Universitas Negeri Semarang. He is a member of the Association of Indonesian Educational Technology Study Programs (APSTPI) and the Indonesian Curriculum Developers Association (HIPKIN). His research interests are in the field of program evaluation and educational management.

**Nurussa'adah** is a lecturer in the Department of Curriculum and Educational Technology, Universitas Negeri Semarang. Obtained a bachelor's degree in the Didactic Curriculum of Universitas Negeri Semarang, graduated with a master's degree in Developmental Psychology at the Universitas Indonesia. She is interested in researching the field of Educational Psychology. Many of his writings were published in journals and seminar proceedings. Currently she is a member of the Association of Indonesian Educational Technology Study Programs (APSTPI).

**Christian.** He is a lecturer at the Department of Curriculum and Educational Technology, Faculty of Education, UNNES, Central Java, Indonesia. Has completed his Masters study in the field of Educational Technology at the 2017, UNIMED, Medan, North Sumatera, Indonesia.