Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5, 892-913
2024
Publisher: Learning Gate
DOI: 10.55214/25768484.v8i5.1792

Local wisdom-based assessment of historical thinking skills: Developing and testing an instrument for junior high school students

Ofianto[1], Rusdi[1], Aisiah[1], Ridho Bayu Yefterson[1], Tri Zahra Ningsih[2*], M Afrillyan Dwi Syahputra[3], Elfa Michellia Karima[1]
[1]Department of History Education, Faculty of Social Science, Universitas Negeri Padang, Padang, Indonesia.
[2]Department of Educational Research and Evaluation, Graduate School, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia; trizahra.2023@student.uny.ac.id (T.Z.N.).
[3]Department of of History Education, Universitas Riau, Indonesia.

Abstract: The aim of this research is to develop an assessment instrument based on local wisdom from the Pagaruyung Palace to measure the historical thinking skills of junior high school students. The research methodology includes the use of a modified ADDIE model, starting with a needs analysis through literature review and interviews with 18 history teachers. The initial instrument was designed and validated by experts in history education, language, and evaluation. The instrument was then tested on 136 eighth-grade students in Solok Regency, and the data were analyzed using the Quest and Parcale programs to evaluate model fit, item difficulty, information function, and Standard Error of Measurement (SEM). The results indicate that the developed instrument has high validity and reliability, with a reliability value of 0.92. All items meet the established criteria and fit the expected measurement model. The integration of local wisdom enhances the relevance and effectiveness of history learning, enriching students' understanding of cultural and historical values. Therefore, the assessment instrument based on local wisdom from the Pagaruyung Palace is effective in measuring the historical thinking skills of junior high school students. This instrument provides a more accurate and contextual assessment of students' abilities to analyze, interpret, and use historical evidence.

Keywords: *Historical thinking, Instrument assessment, Local wisdom, Middle school, Pagaruyung heritage.*

1. Introduction

Historical thinking skills have emerged as a crucial component in modern history education, where these abilities are essential for understanding the complexities of the past and connecting historical events to contemporary contexts (Díez-Bedmar, 2022; Tirado-Olivares et al., 2023; Wilke et al., 2023) With advances in technology and globalization, historical analysis skills are increasingly needed to explore the meaning of past events and apply them in a modern context. In history education, historical thinking skills involve the analysis, interpretation, and critical evaluation of historical evidence (Bunari et al., 2023; Gestsdóttir et al., 2018; Mohidat, 2019; Ofianto, Aman, Sariyatun, Bunari, et al., 2022; Reisman et al., 2019). These skills not only shape critical citizens but also deepen their knowledge of history (Gracia et al., 2023; López-García, 2023; Vijayakumar & Ahmad, 2023). The development of these skills has become increasingly important in the era of globalization and technological advancement, where students are required to navigate complex historical information and relate it to current issues (Andal & Hermosa, 2023).

Despite their importance, developing historical thinking skills often faces challenges in implementation at the secondary school level. One of the primary challenges is the lack of holistic and contextually relevant assessment instruments. Most existing instruments, such as those developed by

* Correspondence: trizahra.2023@student.uny.ac.id

Seixas et al. (Seixas et al., 2015) and Smith (Smith, 2017), tend to be generic and do not consider local wisdom, limiting their relevance in educational contexts in countries like Indonesia.

In Indonesia, local wisdom plays a vital role in enriching history education. A prominent example is the Pagaruyung Palace in West Sumatra, which serves as a symbol of Minangkabau history and culture. Incorporating local wisdom into history education not only improves students' understanding of historical events, but also strengthens their cultural identity (Amin et al., 2022; Fadli et al., 2021; Pratama et al., 2017; Yefterson et al., 2020). Local wisdom from Pagaruyung Palace, for instance, offers unique insights into the history and culture of Minangkabau, helping students to understand historical events within a richer and more meaningful context (Ofianto, Erniwati, Fitrisia, et al., 2023).

Current assessment instruments, however, have yet to fully utilize the potential of local wisdom in measuring historical thinking skills. For example, Ningsih et al. (Ningsih et al., 2019) employed primary evidence without considering local wisdom, while Ofianto et al. (Ofianto, Aman, Sariyatun, Ningsih, et al., 2022) focused more on analysis through essay tests without integrating local wisdom. These shortcomings highlight the urgent need to develop more contextual and relevant assessment instruments, particularly for assessing the historical thinking skills of students in Indonesia.

This research aims to address this gap by developing an assessment instrument for historical thinking skills based on local wisdom from Pagaruyung Palace. By integrating cultural and historical values into the assessment instrument, this study not only seeks to enhance the effectiveness of history education but also to enrich the representation of local cultural heritage in education. Given these challenges, this research seeks to address the specific problem of how to develop a more contextually relevant assessment instrument that integrates local wisdom from Pagaruyung Palace to effectively measure historical thinking skills among junior high school students in Indonesia. This study aims to fill the gap in existing assessment tools by providing an instrument that reflects both global standards and local cultural heritage.

2. Method

2.1. Research Design

This study utilized a modified ADDIE framework to suit the research needs, beginning with a needs analysis through a literature review and interviews with 18 history teachers to identify issues in assessing students' historical thinking skills. The literature review evaluated various previously developed instruments, revealing a lack of integration of local wisdom in assessment tools. Based on the needs analysis, the initial instrument design incorporated key components of historical thinking skills, such as document classification, artifact identification, text analysis, and linking information from various sources. This instrument aimed to measure students' ability to analyze and use historical evidence with a contextual approach based on the local wisdom of Pagaruyung Palace.

The designed instrument was validated by experts from various fields, including history education, language, and evaluation. The validation process involved initial assessment, discussions to reach consensus, and revisions based on expert feedback, ensuring accuracy in historical content, pedagogical relevance, language clarity, and evaluation reliability. Following validation, the instrument was tested on 136 eighth-grade students from various junior high schools in Solok Regency. The student sample was randomly selected to ensure diverse representation. Data from the trial were analyzed using the Quest and Parcale programs to evaluate model fit, item difficulty level, information function, and Standard Error of Measurement (SEM). The implementation of this trial aimed to obtain representative and valid data to assess the instrument's reliability and validity.

Given the nature of the research, a multiple-choice test format was chosen to evaluate the historical thinking skills of students. While it is acknowledged that multiple-choice questions can sometimes limit the depth of assessment, particularly for complex skills such as analysis, interpretation, and critical thinking, this format was selected for several reasons. First, it allows for efficient data collection and analysis from a large sample size, which is essential for ensuring the reliability and validity of the instrument across diverse student populations. Second, multiple-choice tests provide consistent and

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

reliable scoring, reducing subjectivity and ensuring that all students are evaluated based on the same criteria. Third, the multiple-choice questions were carefully designed to measure specific historical thinking skills, such as document classification and source analysis, thereby maintaining the depth and rigor of the assessment.

However, the research also recognizes the limitations of this format. As such, the results of this study will lay the groundwork for future research that may incorporate more varied assessment formats, such as essay-based or document-based questions (DBQs), to further explore students' historical thinking abilities in a more comprehensive manner.

The stages of the research can be seen in the diagram below.

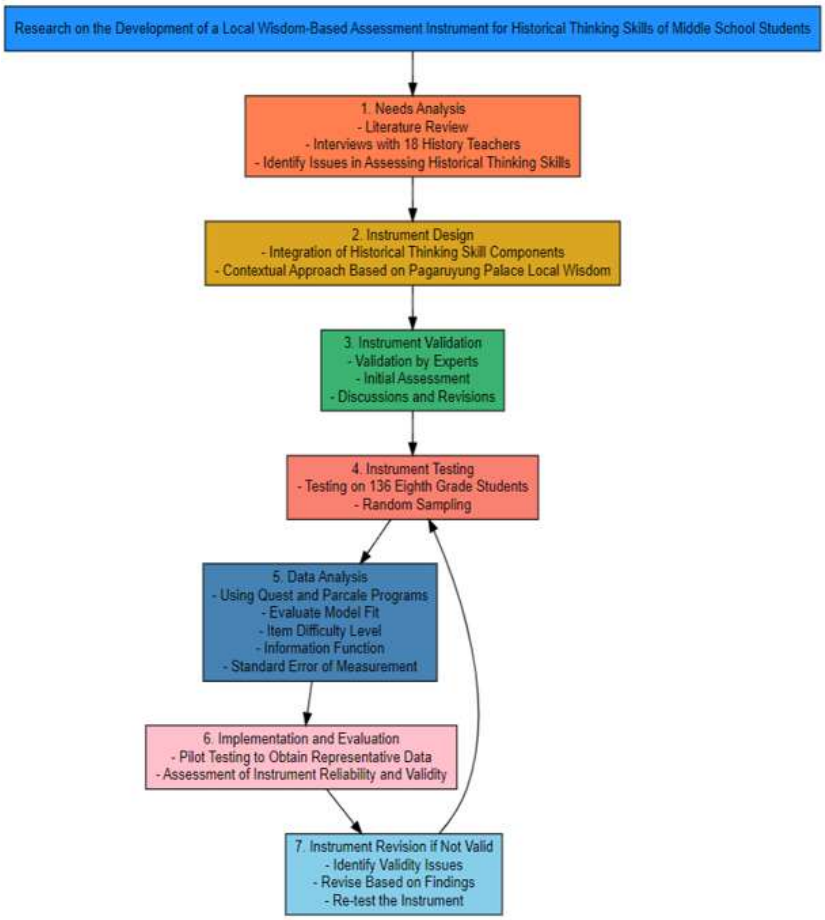

Figure 1.
Research design.

2.2. Participant

This research involved a population of junior high school students in Solok Regency, with a population of 7304 students. For the research sample, three classes at SMP Negeri 1 Solok were randomly selected, with a total sample of 136 students. SMP Negeri 1 Solok was chosen specifically due to its reputation as one of the leading schools in the region, known for its high academic standards and active participation in cultural and historical education initiatives. The selection of three classes as samples was carried out with consideration of sufficient representation of the population, as well as allowing sufficient data to be collected for analysis. Sample selection used the Raosoft.com sample size

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

calculator application with calculations of 5% margin error, 95% confidence level, 92 population, and 10% response distribution, so that the sample needed to declare this research valid was 136 students.

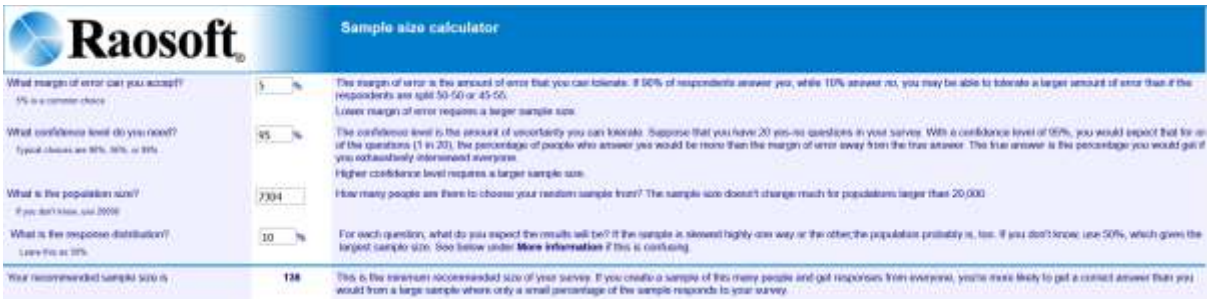

Figure 2.
Sample size calculator.

2.3. Instruments and Data Collection

In this research method, instruments and data collection techniques are carried out through several carefully designed stages. First, data collection related to identifying research problems was carried out through literature analysis and observational studies. Second, instrument validation is carried out through validation sheets by experts. This validation sheet is used to comprehensively measure the validity and reliability of the assessment instrument, as well as to evaluate the suitability of the instrument with the research objectives and conceptual framework that have been previously established. Third, instrument trial data was collected through tests given to 136 junior high school students in Solok Regency. The test given is a multiple-choice test specifically designed to test students' understanding of historical material based on primary sources, as well as to evaluate their ability to analyze, interpret and conclude historical information. The choice of a multiple-choice format, despite its limitations in assessing deeper analytical skills, was made to ensure efficient data processing, consistent evaluation, and the ability to effectively measure specific historical thinking skills relevant to the study's objectives.

2.4. Data Analysis Technique

The data obtained were analyzed using the Quest program to verify the suitability of the model and the level of difficulty index, as well as the Parcale program to obtain information functions and Standard Error of Measurement (SEM). The data in this study are polytomous data from three categories. The content validity test was carried out by experts and the data was analyzed using the Aiken formula. Data analysis techniques include several aspects, namely: conformity with PCM, item validity, item difficulty level, instrument reliability, information function, and Standard Error of Measurement (SEM).

Compliance with the PCM was tested based on the mean and standard deviation of the INFIT MNSQ. If the mean INFIT MNSQ value is close to 1.00 and the standard deviation is close to 0.00, then the entire item is considered to fit in the PCM 1 PL model. This analysis was carried out with the help of the QUEST program. The validity of the items was evaluated based on the MNSQ INFIT value range between 0.77 and 1.30, in accordance with Adams & Khoo [22], which indicates that the items are in good condition or suitable for use. Instrument reliability is estimated from the QUEST program analysis output for item reliability and test participant reliability. According to Mardapi [23], an instrument is considered good enough if it has a reliability coefficient equal to or greater than 0.70.

The item difficulty index is estimated using the QUEST program, with an item considered good if its difficulty index is between -2 and +2. The information function and SEM are estimated with the help of the Parcale program. Based on this information and SEM function, a test or instrument can be said to be suitable for testing on test participants with certain abilities (θ), whether high, low or medium.

Through this comprehensive analysis, the instruments used in the research are expected to be able to provide results that are valid, reliable and in accordance with the established model.

3. Results

This research aims to develop an evaluation instrument based on local wisdom to measure the historical thinking skills of junior high school students through the use of primary evidence from the Pagaruyung Palace. In the context of history education in Indonesia, there is an urgent need to improve students' historical analysis skills, especially in relating past events to the context of their time. The instrument developed is expected to provide a more accurate measurement of students' abilities in analyzing, interpreting and using historical evidence.

This research methodology uses the ADDIE (Analyze, Design, Development, Implementation, and Evaluation) framework which is modified according to research needs. The needs analysis was carried out through a literature review and observations involving interviews with 18 history teachers to identify assessment problems and the instruments used. Based on this needs analysis, a preliminary design of the instrument was developed and validated by experts from various fields, including history, education, language, and evaluation.

Instrument testing was carried out on 136 class VIII students from various junior high schools in Solok Regency. Data from this trial were analyzed using the Quest and Parcale programs to check model fit, item difficulty level, information function, and Standard Error of Measurement (SEM). The results of this analysis determine whether the instrument developed is in accordance with the expected model and can be used for wider implementation.

3.1. Need Analysis

3.1.1. Results of the Library Study

Literature review is an important step in this research to identify and analyze similar research that has been conducted previously. Through in-depth literature analysis, we evaluate various historical thinking skills assessment instruments that have been developed and implemented in various educational contexts. Some relevant research includes the work of Ningsih et al. (Ningsih et al., 2019), Ofianto et al.(Ofianto, Aman, Sariyatun, Ningsih, et al., 2022), Seixas et al. (Seixas et al., 2015), and Smith (Smith, 2017).

These studies provide valuable insight into the methods and approaches used to assess students' historical thinking skills. However, we identified shortcomings in terms of integrating local wisdom as a basis for assessment. Existing instruments tend not to take relevant local cultural context into account.

Seixas et al. (Seixas et al., 2015) used a test to assess historical thinking skills, but this instrument is not based on local wisdom and is only relevant to the educational context in Canada. Smith (Smith, 2017) assessed all aspects of historical thinking skills using a multiple-choice approach, but also did not integrate local wisdom and was limited to the United States context. Ningsih et al. (Ningsih et al., 2019) uses primary evidence to assess historical thinking skills, but their assessment model refers to Seixas et al. and does not pay attention to aspects of local wisdom. Ofianto et al. (Ofianto, Aman, Sariyatun, Ningsih, et al., 2022) focuses on the analyze and consequence aspect through essay tests, but does not integrate local wisdom.

This research seeks to fill this gap by developing an evaluation instrument based on local wisdom from the Pagaruyung Palace, which is expected to make a significant contribution in measuring students' historical thinking skills in their local cultural context. This integration of local wisdom is a major innovation that is expected to increase the relevance and effectiveness of assessment instruments in history education in Indonesia.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

Table 1.
Literature review analysis.

Researcher	Year	Equality	Difference
Seixas et al.	2015	Using tests to assess historical thinking skills	Not based on local wisdom, only relevant to the Canadian context
Smith	2017	Assess all aspects of historical thinking skills	Not based on local wisdom, only relevant to the United States context
Ningsih et al.	2019	Using primary evidence to assess historical thinking skills	Not based on local wisdom, referring to the model developed by Seixas et al.
Ofianthus et al.	2022	Using essay tests to assess historical thinking skills	Not based on local wisdom, only assessing analysis and consequence aspects
This research	2024	Uses primary evidence and essay tests to assess historical thinking skills	Based on local wisdom, relevant to the local context, especially the use of primary evidence from the Pagaruyung Palace

3.1.2. Interview Result

As an integral part of the needs analysis in this research, in-depth interviews were conducted with 18 history teachers to identify the main problems faced in the assessment and use of instruments for evaluating students' skills in using historical sources.

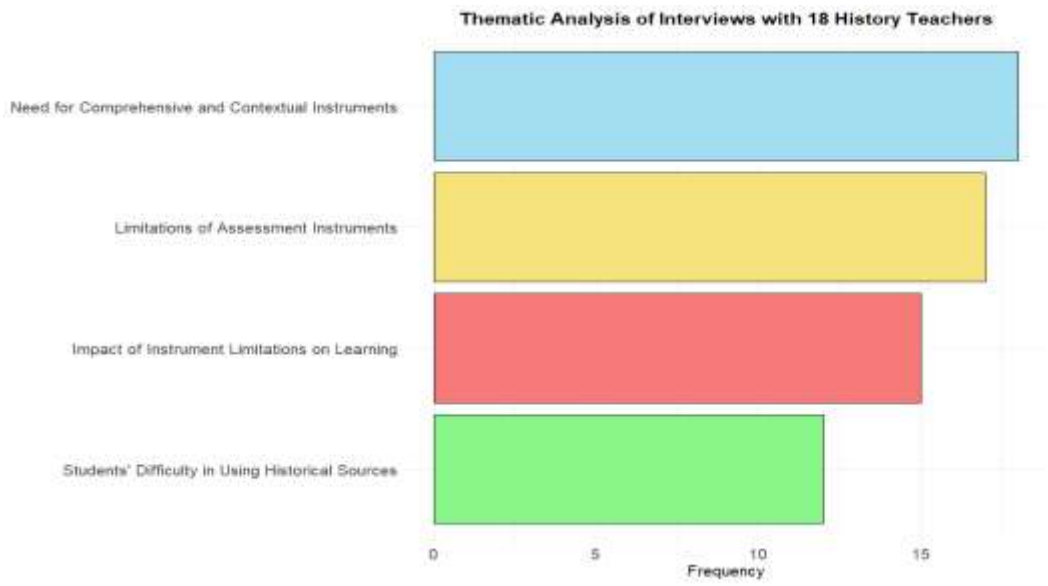

Figure 3.
Result of interview analysis.

From the results of the interviews, it was revealed that teachers experienced significant challenges in assessing students' historical thinking skills. The main problems identified include the limitations of existing assessment instruments, which are often unable to measure students' skills holistically. Existing assessment instruments tend to focus on testing factual knowledge about historical events rather than students' critical analysis and interpretation of historical evidence.

Teachers also report that students often have difficulty using historical sources effectively. This difficulty is caused by the lack of instruments that can help students develop skills in correctly

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

interpreting, analyzing and evaluating historical evidence. In addition, the conventional instruments currently used do not adequately support accurate and comprehensive assessments, so assessment results often do not reflect students' overall abilities. The limitations of existing assessment instruments not only have an impact on measuring students' skills, but also affect the history learning process itself.

Teachers expressed that more comprehensive and contextual instruments are needed to help students develop better historical thinking skills that are relevant to their cultural context. Therefore, this research seeks to develop an assessment instrument that is more comprehensive and based on local wisdom. It is hoped that this instrument can provide a more accurate and relevant evaluation of students' historical thinking skills, especially in the context of using primary evidence from the Pagaruyung Palace. Thus, this instrument will not only improve the quality of assessment, but will also enrich the history learning process by including significant elements of local wisdom.

The development of this instrument aims to meet the need for a more holistic and contextual assessment, as well as supporting the development of students' critical analysis skills in studying history. Thus, this research contributes to improving the quality of history education in Indonesia, as well as maintaining and developing local cultural heritage through a more integrated and relevant approach.

3.2. Design Stage

As part of the development stage, we designed an initial evaluation instrument grid to measure the historical thinking skills of junior high school students based on local wisdom from Pagaruyung Palace. This grid was designed by considering various important components that are relevant to the research objectives, namely to measure students' abilities in classifying, identifying, analyzing and linking information from various historical sources.

The initial grid includes various indicators of historical thinking skills, such as the ability to classify historical documents, identify artifacts, analyze historical texts, and relate information from multiple sources to form a coherent narrative. Each component in this grid is designed to ensure that the instrument developed meets the research objectives, namely to provide an accurate and comprehensive assessment of students' historical thinking skills.

Compliance with the research objectives is achieved through an emphasis on the use of primary evidence from the Pagaruyung Palace, allowing students to develop deeper skills of critical analysis and historical interpretation. By integrating local wisdom, this instrument is not only culturally relevant but also enhances students' understanding of their own historical context. The instrument grid is presented in Table 2.

Table 2.
Historical thinking skills assessment instrument grid.

No.	Indicator	Cognitive level	Question	Answer options
1.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	Types of historical documents that may contain information about Pagaruyung Palace are….	A. Letter from the king to the royal officials B. An architect's memoirs about the construction of the palace C. Diary of a merchant visiting the palace D. Official government report on celebrations at the palace
2.	Identify artifacts such as architecture, weapons, or household equipment that are relevant to a	C5	Artifacts relevant to the time period during which Pagaruyung	A. Ceramic plates with Hindu-Buddhist motifs B. Arrows and bows used in the defense of the castle

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
	particular time period to understand aspects of life in the past.		Palace was built may include….	C. Stationery from the 14th century D. Modern weapons such as rifles
3.	Analyze historical texts to identify the facts, opinions, and historical context contained therein.	C4	What can be identified from historical texts about Pagaruyung Palace….	A. A historian's opinion on palace architecture B. Facts about the years of construction and important events at the palace C. Historical context about trade around the palace D. A pilgrim's testimony about a spiritual experience at the palace
4.	Analyze artifacts to identify information about their function, materials, and historical context.	C5	How can we find out the function and building materials of Pagaruyung Palace….	A. From an architect's diary B. From trade letters from local traders C. From the analysis of the building structure and its materials D. From the memoirs of a musician about life at court
5.	Relate information from multiple written sources to create a coherent and comprehensive narrative about a historical event.	C5	How can we create a coherent narrative about the history of Pagaruyung Palace….	A. By using only one written source B. By combining information from various written sources C. Just quoting the opinion of a historian D. By ignoring information that doesn't fit our narrative
6.	Linking information from relevant artifacts to create an in-depth and contextual narrative about a time period or historical event.	C6	To create an in-depth narrative about Pagaruyung Palace, we must….	A. Only considers information from one artifact source B. Ignore irrelevant artifacts C. Linking information from various artifacts to provide broader context D. Making assumptions without testing the validity of the information
7.	Identify biases, gaps, or deficiencies in a historical document and evaluate their impact on the understanding of a historical event.	C5	How can we evaluate biases or gaps in historical documents about Pagaruyung Palace….	A. By only trusting what is written in official government documents B. By comparing information from different documents C. By ignoring other people's opinions about the document D. By not questioning the authenticity of the document
8.	Assessing the extent to	C6	What artifact is	A. Ceramic plates with Royal

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
	which particular artifacts can provide significant insight into aspects of life in the past that are relevant to the historical context being studied.		most likely to provide significant insight into daily life at Pagaruyung Palace….	motifs B. Weapons of war used in palace defense C. Gold bracelets worn by the Royal family D. The king's diaries are well preserved
9.	Create alternative narratives or new theories based on a deep understanding of written sources, which can provide new insights or points of view that have not been considered before.	C6	How can we create an alternative narrative about the history of Pagaruyung Palace….	A. By only considering one point of view B. By ignoring conflicting evidence C. By considering information from various sources and points of view D. By making assumptions without being based on existing facts
10.	Developing new hypotheses or innovative interpretive models from existing artifacts, which can help in better understanding a relevant historical event or aspect of life in the past.	C6	Developing a new hypothesis from the artifacts at Pagaruyung Palace can help us….	A. Understand better the daily life of the king and queen in the palace B. Explains why the palace was built in that place C. Determine the selling price of these artifacts on the art market D. Create a fictional story about life in the palace
11.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	Historical documents such as letters from kings to government officials can provide information about….	A. Household appliances used in the palace B. Trade travel between Kingdoms C. Government policy regarding palace construction D. Tradition of religious ceremonies carried out in the palace
12.	Identify artifacts such as architecture, weapons, or household equipment that are relevant to a particular time period to understand aspects of life in the past.	C5	Artifacts relevant to the time period during which Pagaruyung Palace was built may include….	A. Ancient coins used as currency in the Kingdom B. Marine navigation maps from that time C. Oil paintings about daily life in the palace D. War horses used by government forces
13.	Analyze artifacts to identify information about their function, materials, and historical context.	C5	What information can be found from the structural analysis of the	A. Agricultural equipment used around the castle B. Last construction and renovation date C. The exact position of important

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
			Pagaruyung Palace building….	rooms in the palace Reconstruction of the original wall paint color
14.	Identify biases, gaps, or deficiencies in a historical document and evaluate their impact on the understanding of a historical event.	C5	How to evaluate the impact of bias in a diary about Pagaruyung Palace….	A. By comparing daily records with other daily records from the same time B. By only trusting information that fits our own point of view C. By ignoring information that may conflict with existing narratives D. By only considering testimony from people who had a close relationship with the owner of the diary
15.	Assessing the extent to which particular artifacts can provide significant insight into aspects of life in the past that are relevant to the historical context being studied.	C6	Artifacts that may provide significant insight into daily life at Pagaruyung Palace include….	A. Statues of Hindu gods used in religious ceremonies B. Modern watches and jewelry that are not related to the palace's past C. Mural paintings about the history of the kingdom found on the walls of the palace D. Traditional cooking utensils found in palace kitchens
16.	Create alternative narratives or new theories based on a deep understanding of written sources, which can provide new insights or points of view that have not been considered before.	C6	How can we create an alternative narrative about the history of Pagaruyung Palace….	A. By ignoring information from sources that do not fit the main narrative B. By considering different points of view from existing sources C. By only selecting information that supports our own views D. By rejecting information that might challenge existing views
17.	Developing new hypotheses or innovative interpretive models from existing artifacts, which can help in better understanding a relevant historical event or aspect of life in the past.	C6	Developing new interpretive models of Pagaruyung Palace artifacts can help us….	A. Ignoring the old history of the palace B. Better understand life in the castle in the past C. Ignore irrelevant artifacts D. Create a fictional story about life in the palace
18.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to	C4	Information about celebrations and ceremonies at Pagaruyung	A. Diary of an architectural expert B. Trade letters between neighboring kingdoms C. Memoirs of a court musician D. Official government report on

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
	understand the context of a historical event.		Palace may be contained in….	palace activities
19.	Relate information from multiple written sources to create a coherent and comprehensive narrative about a historical event.	C5	Why is it important to compare information from various written sources in creating a historical narrative about Pagaruyung Palace….	A. So that the narrative becomes more complex and difficult to understand B. So that we can find consistency in various sources C. So that we can reject sources that may be inaccurate D. So that we can choose one source that is most beneficial to our own point of view
20.	Assessing the extent to which particular artifacts can provide significant insight into aspects of life in the past that are relevant to the historical context being studied.	C6	How can we assess the relevance and significance of artifacts in the historical context of Pagaruyung Palace….	A. By paying attention only to the most visually interesting artifacts B. By comparing artifacts with similar modern items C. By considering the historical context and function of these artifacts D. By choosing the most expensive and financially valuable artifacts
21.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What is the benefit of examining trade letters between neighboring kingdoms to understand the history of Istana Pagaruyung….	A. To know the types of plants grown around the palace B. To understand political and economic relations between regions C. To find out details of the architecture and interior design of the palace D. To assess the success of religious ceremonies at the palace
22.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What can we learn from a court musician's memoir about life at Pagaruyung Palace….	A. The daily life of the king and queen in the palace B. Defense strategy used to protect the castle C. Trade policies implemented by the Royal government D. Infrastructure development around the palace
23.	Relate information from multiple written sources to create a coherent and comprehensive narrative about a historical event.	C5	Why is it important to consider different points of view in creating a	A. So that the narrative becomes simpler and easier to understand B. So that we can strengthen our own arguments C. So that we can understand

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
			historical narrative about Pagaruyung Palace….	different perspectives on an event D. So that we can reject points of view that are contrary to our own opinion
24.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	Where can we find information about religious ceremonies held at Pagaruyung Palace….	A. From the memoirs of a merchant who often interacted with the Royal family B. From official government reports about religious activities in the palace C. From the analysis of the structure of palace buildings and the location of places of worship D. From trade letters recording purchases of religious ceremonial equipment
25.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What can we learn from the king's well-preserved diaries….	A. Daily activities of royal family members in the palace B. A war strategy used to protect the kingdom from enemy invasion C. Diplomatic relations with neighboring governments Infrastructure development plans around the palace
26.	Identify biases, gaps, or deficiencies in a historical document and evaluate their impact on the understanding of a historical event.	C5	How can we assess the validity and reliability of the information in a historical memoir about Pagaruyung Palace….	A. By only trusting information that matches our own opinion B. By comparing information with other existing sources C. By ignoring memoirs if the author does not have an academic degree D. By only considering emotionally appealing information
27.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What are the benefits of checking official government reports about the celebrations at Pagaruyung Palace….	A. To find out the number of visitors to the palace each year B. To understand royal traditions and culture that are maintained through ceremonies C. To determine the selling price of artifacts related to the celebration D. To assess the quality of entertainment and food served at the celebration
28.	Identify biases, gaps, or	C5	How can we	A. By only believing testimony

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

No.	Indicator	Cognitive level	Question	Answer options
	deficiencies in a historical document and evaluate their impact on the understanding of a historical event.		consider the testimony of people who were closely related to the owner of the king's diaries….	that supports our own views B. By rejecting testimony that may not fit the existing narrative C. By checking the truth of the testimony through other available evidence D. By not considering testimony that is not relevant to our focus
29.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What are the benefits of examining marine navigation charts from that time to understand the history of Pagaruyung Palace….	A. To find out the pattern of trade travel between Kingdoms B. To understand the weather conditions around the palace area C. To identify the best location to build a defensive fort D. To assess the navigational quality of the ships used at that time
30.	Classify various types of historical documents such as letters, memoirs, diaries, and official government reports to understand the context of a historical event.	C4	What can we learn from the ancient coins used as currency in the Pagaruyung kingdom….	A. The currency exchange rate at that time B. The kingdom's trade journey with neighboring kingdoms C. The banking system implemented by the government D. The artistic and artistic value represented by the design of the coin

3.3. Development

3.3.1. Expert Validation

To ensure that the assessment instrument developed has high validity, the validation process involves various experts from several related fields. The experts involved in this validation include historians, education experts, linguists and evaluation experts. Historians provide deep insight into the accuracy of historical content and the relevance of artifacts and documents used in assessment instruments. Educational experts help ensure that assessment instruments comply with pedagogical principles and can be used effectively in learning contexts. Linguists check the clarity, accuracy, and readability of the language used in the instrument, ensuring that the instructions and questions are easily understood by students. Meanwhile, evaluation experts provide input regarding the reliability and validity of the instruments, as well as ensuring that the assessment methods used are appropriate and in line with the research objectives.

The validation process is carried out through several stages to ensure that the instrument developed meets high standards. This stage includes initial assessment by experts, discussion to reach consensus, and revision of the instrument based on input from experts. Each expert assessed the instruments based on predefined criteria, such as content accuracy, pedagogical relevance, language clarity, and evaluation reliability. After the initial assessment, a discussion is held to reach consensus regarding necessary improvements and adjustments. These discussions allow experts to share views and provide recommendations based on their respective expertise. Based on input and recommendations from experts, revisions were made to the instrument to improve its quality and validity.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

Feedback provided by experts is invaluable in the instrument development process. Some of the points of feedback and revisions made include adjusting historical content, improving readability and clarity, increasing pedagogical relevance, and strengthening reliability and validity. Historians provide suggestions to improve or add more accurate and relevant information related to the historical context of Pagaruyung Palace. Linguists suggested improvements to word usage and sentence structure to ensure the instrument was easy for students to understand. Education experts provide recommendations to ensure that the instrument can be used effectively in the learning process, including adjustments in the way it is presented and the types of questions. Evaluation experts provide input to ensure that the instrument can produce consistent and reliable assessments, including recommendations for further testing and statistical analysis.

By going through this comprehensive validation process, the assessment instruments developed are expected to meet high-quality standards and provide accurate and relevant evaluations of middle school students' historical thinking skills.

Table 3.
Average score of expert assessment.

Validated aspects	Assessment indicators	Mean	Recommendation
Historical content accuracy	Historical accuracy Relevance of historical context to the topic Compatibility with historical sources Completeness of historical information Consistency of historical narrative	4.5	Continuously update and verify historical facts used in the instrument. Add more recent and relevant historical sources.
Pedagogical relevance	Compliance with pedagogical principles Effectiveness in learning Applicability in the curriculum Suitability to student needs Flexibility of instrument use	4.2	Continuously evaluate the effectiveness of instruments in learning. Wider field trials to adapt instruments to diverse student needs.
Language clarity and readability	Language Clarity Text readability Correct terminology Simplicity of sentences Suitability of language style to students' level of understanding	4.0	Revise the text to ensure that the language used is simple and easy for students to understand. Involve linguists in every stage of instrument development.
Reliability and validity of evaluation	Consistency of assessment Instrument reliability Validity of assessment results Accuracy in measuring historical thinking skills Relevance to research objectives	4.3	Perform further testing and statistical analysis to ensure the reliability and validity of the instrument. Strengthen instruments with more comprehensive evaluation methods.

The validation table for the historical thinking skills assessment instrument includes four main aspects that are validated, namely historical content accuracy, pedagogical relevance, language clarity and readability, and evaluation reliability and validity. The accuracy of historical content includes accuracy of historical facts, relevance of historical context to the topic, conformity with historical sources, completeness of historical information, and consistency of historical narrative. The validation

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

results show an average value of 4.5, which indicates that the historical content presented is quite accurate. However, it is recommended to continuously update and verify the historical facts used in the instrument. The addition of more up-to-date and relevant historical sources is also recommended to maintain the quality of historical content. Pedagogical relevance includes conformity with pedagogical principles, effectiveness in learning, applicability in the curriculum, suitability to student needs, and flexibility in the use of instruments. With an average score of 4.2, these results indicate that the instrument is quite relevant pedagogically.

The recommendation given is to carry out continuous evaluation of the effectiveness of the instrument in learning and conduct wider field trials to adapt the instrument to the needs of diverse students. Language clarity and readability include clarity of language, readability of text, accuracy of terminology, simplicity of sentences, and suitability of language style to the level of student understanding. With an average score of 4.0, this aspect requires special attention in text revision to ensure that the language used is simple and easy for students to understand. It is recommended to involve language experts at every stage of instrument development to improve clarity and readability. Evaluation reliability and validity include assessment consistency, instrument reliability, validity of assessment results, accuracy in measuring historical thinking skills, and relevance to research objectives. With an average value of 4.3, these results indicate that the instrument is quite reliable and valid. However, it is recommended to carry out further testing and statistical analysis to ensure the reliability and validity of the instrument. In addition, strengthening instruments with more comprehensive evaluation methods will help improve the quality of assessments. Overall, these validation tables provide important guidance for refining and refining historical thinking skills assessment instruments, with a focus on improving content accuracy, pedagogical relevance, language clarity, and evaluation reliability and validity.

3.3.2. Reliability Test Results

In an effort to ensure the reliability and validity of the historical thinking skills assessment instrument, a comprehensive statistical analysis of various reliability metrics was conducted. The results of the reliability test are presented in the table below:

Table 4.
Reliability test results.

Metric	Value
Raw alpha	0.93
Standardized alpha	0.93
Guttman lambda 6	0.95
Average inter-item correlation (Average_r)	0.32
Signal-to-noise ratio (S/N)	14
Standard error of alpha (ase)	0.0083
Mean	0.45
Standard deviation (sd)	0.29
Median inter-item correlation (Median_r)	0.31

The results of the reliability analysis show that the instrument used has a very high level of internal consistency, as evidenced by the Cronbach's Alpha value of 0.93. This value indicates that the instrument has an excellent ability to produce consistent results when used in a variety of similar situations. In addition, the standardized Alpha value is also 0.93, which strengthens the finding that the reliability of the instrument remains high after the items are standardized.

The Guttman Lambda 6 value of 0.95 further confirms the high reliability of this instrument, indicating that the variance explained by the common factor is very high, thereby increasing confidence in the internal consistency of the instrument. The average correlation between items (average_r) of 0.32

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

indicates that the items in this instrument have a fairly good correlation with each other, which is a positive indication for internal consistency. A signal to noise ratio (S/N) of 14 indicates that the variability produced by the signal (item consistency) is much higher than the noise (error), which supports high reliability.

The low standard error of the Alpha (ase) estimate of 0.0083 indicates that the Alpha estimate is quite accurate and reliable. The mean item score was 0.45 with a standard deviation of 0.29, providing a uniform distribution of item scores. The median inter-item correlation of 0.31 is consistent with the average inter-item correlation, indicating that the majority of items correlate well with each other.

The 95% confidence range for Alpha is between 0.92 to 0.95 according to Feldt and Duhachek's method, indicating that this high reliability value is stable and trustworthy. This indicates that this instrument has very good consistency in measuring the intended construct, and can be relied upon for use in further research or evaluation.

Overall, this instrument shows excellent performance in terms of reliability and internal consistency, making it a valid and reliable tool for data collection in the context of educational research or other evaluations. These results support the validity of the instrument as a measuring tool capable of providing consistent and accurate results in various applications.

3.3.3. Limited Trial

To ensure the validity and reliability of the assessment instrument developed, a limited trial was conducted involving 136 class VIII students. This sample of students was randomly selected from various schools in Solok Regency, both public and private schools, to obtain wider and more diverse representation. The sample consisted of 136 students who were in class VIII, with a demographic distribution that included various social and economic backgrounds. The average age of students is 13-14 years, with an equal number of male and female students. Additionally, there are variations in ethnic and cultural backgrounds that provide a more comprehensive picture of the student population in Solok Regency. The schools participating in this trial have diverse characteristics. There are public and private schools with different facilities, from schools in city centers to those in rural areas. Additionally, the participation of schools with varying academic backgrounds allows for a more thorough assessment of the instrument's effectiveness in a variety of educational contexts.

By involving a diverse sample, this limited trial is expected to provide representative and valid data to evaluate the reliability and validity of the instrument for assessing students' historical thinking skills. The results of this trial will be used to make further adjustments and improvements to the instrument being developed, so that it can be used effectively in the context of history learning at the junior high school level.

Table 5.
Difficulty level of question details.

Question number	Degree of difficulty	Information
1	-1.5	Meet the criteria
2	-1.0	Meet the criteria
3	-0.8	Meet the criteria
4	-1.2	Meet the criteria
5	-1.4	Meet the criteria
6	-1.7	Meet the criteria
7	-1.0	Meet the criteria
8	-1.5	Meet the criteria
9	-1.9	Meet the criteria
10	-2.0	Meet the criteria
11	-1.3	Meet the criteria
12	-1.2	Meet the criteria

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

Question number	Degree of difficulty	Information
13	-1.6	Meet the criteria
14	-1.0	Meet the criteria
15	-1.5	Meet the criteria
16	-1.9	Meet the criteria
17	-1.8	Meet the criteria
18	-1.3	Meet the criteria
19	-1.4	Meet the criteria
20	-1.6	Meet the criteria
21	-1.4	Meet the criteria
22	-1.5	Meet the criteria
23	-1.2	Meet the criteria
24	-1.3	Meet the criteria
25	-1.5	Meet the criteria
26	-1.0	Meet the criteria
27	-1.2	Meet the criteria
28	-1.3	Meet the criteria
29	-1.2	Meet the criteria
30	-1.4	Meet the criteria

The table above summarizes the results of item analysis from the historical thinking skills assessment instrument for junior high school students based on level of difficulty and information. All items meet the criteria of a difficulty range between -2.00 to +2.00, indicating that this instrument can measure students' historical thinking skills well. Furthermore, to test the model fit, an INFIT MNSQ test was conducted. The test results are presented in Table 6.

Table 6.
Model fit.

Item Number	INFIT MNSQ	Information
Item 1	1.02	Fits the model
Item 2	0.98	Fits the model
Item 3	1.01	Fits the model
Item 4	0.99	Fits the model
Item 5	1.03	Fits the model
Item 6	0.97	Fits the model
Item 7	1.00	Fits the model
Item 8	1.01	Fits the model
Item 9	0.99	Fits the model
Item 10	1.03	Fits the model
Item 11	0.98	Fits the model
Item 12	1.00	Fits the model
Item 13	1.02	Fits the model
Item 14	0.97	Fits the model
Item 15	1.01	Fits the model
Item 16	0.99	Fits the model
Item 17	1.00	Fits the model
Item 18	1.02	Fits the model
Item 19	0.98	Fits the model

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

Item Number	INFIT MNSQ	Information
Item 20	1.00	Fits the model
Item 21	1.03	Fits the model
Item 22	0.99	Fits the model
Item 23	1.01	Fits the model
Item 24	1.00	Fits the model
Item 25	1.02	Fits the model
Item 26	0.98	Fits the model
Item 27	1.01	Fits the model
Item 28	0.99	Fits the model
Item 29	1.03	Fits the model
Item 30	1.00	Fits the model

The table 6 presents the results of the INFIT MNSQ (Infit Mean Square) test for each item in the instrument, evaluating the fit of each item to the model. The INFIT MNSQ values for all items range between 0.97 and 1.03, indicating that they all fit well within the acceptable range for model fit, typically considered to be between 0.70 and 1.30. Specifically, each item has an INFIT MNSQ value close to 1.00, which demonstrates that the items conform to the expected measurement model and perform consistently in measuring the targeted construct. This consistency across all items supports the validity and reliability of the instrument, confirming that it is well-suited for assessing historical thinking skills in junior high school students. The results show that each item meets the model fit criteria, as indicated by the "Fits the model" information for all items listed. This ensures that the instrument can provide accurate and reliable assessments of students' abilities.

Based on the results of the expanded test data analysis, the characteristics of the competency assessment instrument for measuring historical thinking skills of junior high school students, based on local wisdom from the Pagaruyung Palace, were obtained. The estimation results of 30 items were conducted with 136 students at a probability level of 0.5 and using polytomous scoring according to PCM 1 PL of the three categories, the results are shown in Table 7.

Table 7.
Estimation of historical thinking skills item instrument.

Description	Trial
Testee reliability	0.92
Value of INFIT MNSQ	1.00
RMSEA	0.88 – 1.15
Mean and standard deviation of INFIT MNSQ	1.01 ± 0.05
Standard error of measurement (SEM)	0.04 – 0.06

Table 7 presents the estimation results from the instrument for assessing students' historical thinking skills. With a reliability value of 0.92, this instrument shows good consistency. The average INFIT MNSQ value is 1.00 with a standard deviation of 0.05 indicating that the items in this instrument are in accordance with the PCM model. The RMSEA range from 0.88 to 1.15 also indicates a good fit, while the SEM indicates that the level of measurement error is within acceptable limits.

4. Discussion

This research aims to develop an evaluation instrument based on local wisdom to measure the historical thinking skills of junior high school students through the use of primary evidence from the Pagaruyung Palace. In the context of history education in Indonesia, there is a growing urgency to enhance students' historical analysis skills, particularly in their ability to relate past events to the

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

context of their time. In the context of history education in Indonesia, there is an urgent need to improve students' historical analysis skills, especially in relating past events to the context of their time. The instrument developed is expected to provide a more accurate measurement of students' abilities in analyzing, interpreting and using historical evidence. This research builds on previous studies by addressing existing gaps and limitations. This research supports previous research by completing existing shortcomings. Several previous studies such as Seixas et al. (Seixas et al., 2015) and Smith (Smith, 2017) developed tests to assess historical thinking skills but are not based on local wisdom and are only relevant in the context of Canada and the United States. Ningsih et al. (Ningsih et al., 2019) use primary evidence but does not consider local wisdom. Ofianto et al. (Ofianto, Aman, Sariyatun, Ningsih, et al., 2022) focuse on analysis through essay tests without integration of local knowledge. This research seeks to fill this gap by integrating local wisdom from the Pagaruyung Palace, which is expected to increase the relevance and effectiveness of assessment instruments in history education in Indonesia.

The research results show that the instrument developed has high reliability and validity. Specifically, the reliability test indicates that the instrument possesses a very high level of internal consistency, evidenced by a Cronbach's Alpha value of 0.93. All items in the instrument meet the criteria of a difficulty range between -2.00 to +2.00, indicating that this instrument can measure students' historical thinking skills well. One of the main advantages of this research lies in the integration of local wisdom and the use of primary evidence, which collectively provide a more culturally relevant context for students (Anderberg et al., 2018; Fariyatul & Bandono, 2017; Pellegrini & Vivanet, 2021; Tally & Goldenberg, 2005). The use of primary evidence allows students to feel the "atmosphere" of the time period studied, connect themselves to the past, and form a holistic understanding of the historical context (Ningsih et al., 2019; Ofianto, Erniwati, Fitrisia, et al., 2023; Ofianto, Erniwati, & Fitrisia, 2023). Apart from that, this research also emphasizes comprehensive validation and testing, involving experts from various fields and representative student samples.

The local wisdom-based evaluation instrument developed in this research has several significant advantages. First, the integration of local wisdom from the Pagaruyung Palace provides higher cultural relevance for students, allowing them to understand history in their own cultural context. This helps students develop more contextual and in-depth historical thinking skills (Muhammad et al., 2022; Priamantono, 2020). Second, the use of primary evidence in this instrument allows students to interact directly with authentic historical sources, which not only enriches the learning experience but also enhances their critical analysis and interpretation skills (Eka Yuliana Rahman et al., 2023; Pajriah & Suryana, 2021). Third, this instrument is designed to measure various aspects of historical thinking skills, such as document classification, artifact identification, and historical text analysis, thereby providing a more comprehensive assessment of students' abilities (Salbella & Kumalasari, 2020). Fourth, this instrument has gone through a rigorous validation process by experts from various fields, ensuring the accuracy of the content, pedagogical relevance, clarity of language, as well as the reliability and validity of the evaluation. Fifth, this instrument was tested on a representative sample of junior high school students, ensuring that the results obtained are reliable and widely applicable (Ufie et al., 2020). Sixth, the results of the reliability test show that this instrument has a very high level of internal consistency, which means that this instrument can produce consistent results when used in various similar situations (Pratama et al., 2017). Finally, this instrument not only functions as an evaluation tool but also as a means to improve history learning.

By integrating elements of local wisdom, this instrument encourages students to engage more deeply in learning history, understand their cultural heritage, and develop a sense of pride in their local identity (Abidin et al., 2020; Hanes & Stone, 2019; Kleftodimos et al., 2023; Putro & Setyowati, 2022). Thus, this instrument not only provides an accurate and relevant assessment of students' historical thinking skills but also contributes to improving the quality of history education and preserving local cultural heritage. In conclusion, this research has successfully developed an effective evaluation instrument for measuring the historical thinking skills of junior high school students. This instrument shows high reliability and validity, making it a valuable tool in history education in Indonesia.

Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 8, No. 5: 892-913, 2024
DOI: 10.55214/25768484.v8i5.1792

This research has significant implications in the preservation and development of local cultural heritage. In an increasingly globalized world, it is important to maintain and promote local cultural identity as part of humanity's heritage. By encouraging the use of local wisdom in history learning, it is hoped that it can strengthen students' understanding of the cultural and historical values that underlie their society and make a sustainable contribution to the maintenance and development of local cultural heritage while strengthening the sense of pride and identity of local communities in the context of deeper history education. wide.

5. Conclusion

This research aims to develop an evaluation instrument rooted in the local wisdom of Pagaruyung Palace to assess the historical thinking skills of junior high school students. The instrument exhibits high validity and reliability, with a reliability coefficient of 0.92. By integrating local wisdom, the instrument enhances the relevance and effectiveness of history education, fostering a deeper understanding of cultural and historical values. This provides a more accurate and contextually relevant measure of students' ability to analyze, interpret, and apply historical evidence. The findings offer significant practical implications for history education by improving the effectiveness of developing students' historical thinking skills. Furthermore, the inclusion of local wisdom helps preserve and promote local cultural heritage, which is essential in an increasingly globalized world. Consequently, this research makes both academic and practical contributions, reinforcing students' cultural identity and fostering pride in their heritage.

However, the study has several limitations. First, the trial was conducted only in Solok Regency and focused specifically on the local wisdom of Pagaruyung Palace, which may limit the generalizability of the findings to other regions in Indonesia. Second, the sample size was relatively small, consisting only of eighth-grade students from a limited number of schools in one district, potentially limiting the applicability of the results to a broader student population.

Future research should address these limitations by expanding the testing of the developed instrument to various regions of Indonesia, ensuring its validity and reliability across different cultural contexts. This can be achieved by conducting field trials with larger and more diverse samples to obtain more representative data. Additionally, the development of assessment instruments that incorporate local wisdom from different regions, such as traditional practices, customs, and cultural artifacts, will increase the comprehensiveness of the tool. Collaboration with local cultural experts and educators is crucial to ensuring the accuracy and relevance of these instruments. Moreover, more sophisticated analytical methods, such as SEM, multilevel modeling, or longitudinal analysis, should be employed to gain a deeper understanding of the factors that influence students' historical thinking skills. Expanding the validation of similar instruments across different educational levels, such as elementary and high schools, will also provide insights into the progression of historical thinking skills across different stages of student development.

References

[1] Abidin, R., Suryani, N., & Sariyatun, . (2020). Students' Perceptions of 360 Degree Virtual Tour-Based Historical Learning About The Cultural Heritage Area of The Kapitan and Al-Munawar Villages in Palembang City. *International Journal of Social Sciences and Management*. https://doi.org/10.3126/ijssm.v7i3.29764

[2] Amin, A., Alimni, Kurniawan, D. A., Chen, D., & Wirayuda, R. P. (2022). Servation of Bengkulu Local Wisdom: The Application of Syarafal Anam in Preventing Student Radicalism. *International Journal of Instruction*, *15*(3), 931–948. https://doi.org/10.29333/iji.2022.15350a

[3] Andal, R. A. E., & Hermosa, J. P. (2023). Learning Style of Grade Eight Students: Basis on Enhancing Historical Thinking Skills Using Contextualized Learning Activities in Araling Panlipunan. *International Journal of*

Multidisciplinary: Applied Business and Education Research, *4*(5), 1639–1650. https://doi.org/10.11594/ijmaber.04.05.25

[4] Anderberg, L., Katz, R. M., Hayes, S., Stankrauff, A., Hodgetts, M. M., Hurtado, J., Nye, A., & Todd-Diaz, A. (2018). Teaching the teacher: Primary source instruction in American and Canadian archives graduate programs. *American Archivist*, *81*(1), 188–215. https://doi.org/10.17723/0360-9081-81.1.188

[5] Bunari, Fadli, M. R., Fikri, A., Setiawan, J., Fahri, A., & Izzati, I. M. (2023). Understanding history, historical thinking, and historical consciousness, in learning history: An ex post-facto correlation. *International Journal of Evaluation and Research in Education*, *12*(1), 260–267. https://doi.org/10.11591/ijere.v12i1.23633

[6] Díez-Bedmar, M. del C. (2022). Feminism, Intersectionality, and Gender Category: Essential Contributions for Historical Thinking Development. In *Frontiers in Education* (Vol. 7, p. 842580). https://doi.org/10.3389/feduc.2022.842580

[7] Eka Yuliana Rahman, Aldegonda E. Pelealu, Hetreda Terry, Yohanes Burdam, Max Tamon, Meity Najoan, Meike Imbar, Aksilas Dasfordate, & Almen S. Ramaino. (2023). Trends Research On Local Wisdom In History Learning: A Bibliometric Analysis. *Technium Social Sciences Journal*, *49*(1), 272. https://doi.org/10.47577/tssj.v49i1.9814

[8] Fadli, M. R., Sudrajat, A., & Amboro, K. (2021). *The influence of sorogan method in learning history to increase historical understanding and historical awareness*. *10*(1), 300–307. https://doi.org/10.11591/ijere.v10i1.20972

[9] Fariyatul, E., & Bandono, A. (2017). The use of value clarification technique-based- picture story media as an alternative media to value education in primary school. *Harmonia: Journal of Arts Research and Education*, *17*(1), 68. https://doi.org/10.15294/harmonia.v17i1.7469

[10] Gestsdóttir, S. M., van Boxtel, C., & van Drie, J. (2018). Teaching Historical thinking and reasoning: Construction of an observation instrument. *British Educational Research Journal*, *44*(6), 960–981. https://doi.org/10.1002/berj.3471

[11] Gracia, P. R., Morán, B. A., & García-Ceballos, S. (2023). Progression of Historical Thinking in Secondary School Students: Sources and Critical Thinking. *Revista Electronica de Investigacion Educativa*, *25*. https://doi.org/10.24320/redie.2023.25.e09.4338

[12] Hanes, L., & Stone, R. (2019). A model of heritage content to support the design and analysis of video games for history education. *Journal of Computers in Education*, *6*(4), 587–612. https://doi.org/10.1007/s40692-018-0120-2

[13] Kleftodimos, A., Moustaka, M., & Evagelou, A. (2023). Location-Based Augmented Reality for Cultural Heritage Education: Creating Educational, Gamified Location-Based AR Applications for the Prehistoric Lake Settlement of Dispilio. *Digital*, *3*(1), 18–45. https://doi.org/10.3390/digital3010002

[14] López-García, A. (2023). Effectiveness of a teaching methodology based on the theory of historical thinking through active methods and digital resources in Spanish adolescents. In *Frontiers in Education* (Vol. 8, p. 1175123). https://doi.org/10.3389/feduc.2023.1175123

[15] Mohidat, J. M. A.-N. (2019). Effectiveness of Teaching History Using Imaginary Learning Strategy in Improving Historical Thinking Among The Tenth primary Grade students in Both Taibeh and Westieh Brigades in Irbid Governorate. *International Journal of Education and Research*, *7*(5), 151–166.

[16] Muhammad, U. A., Fuad, M., Ariyani, F., & Suyanto, E. (2022). Bibliometric analysis of local wisdom-based learning: Direction for future history education research. *International Journal of Evaluation and Research in Education*, *11*(4), 2209–2222. https://doi.org/10.11591/ijere.v11i4.23547

[17] Ningsih, T. Z., Sariyatun, & Sutimin, L. A. (2019). Development of portfolio assessment to measure the student's skill of using primary source evidence. *New Educational Review*, *56*(2), 101–113. https://doi.org/10.15804/tner.2019.56.2.08

[18] Ofianto, Erniwati, Fitrisia, A., Ningsih, T. Z., & Mulyani, F. F. (2023). Development of Online Local History Learning Media Based on Virtual Field Trips to Enhance the Use of Primary Source Evidence. *European Journal of Educational Research*, *12*(2), 775–793. https://doi.org/https://doi.org/10.12973/eu-jer.12.2.775

[19] Ofianto, O., Aman, A., Sariyatun, S., Bunari, B., Zahra, T. Z. N., & Marni, M. E. P. (2022). Media Timeline Development with the Focusky Application to Improve Chronological Thinking Skills. *International Journal of Learning, Teaching and Educational Research*, *21*(4), 114–133. https://doi.org/10.26803/ijlter.21.4.7

[20] Ofianto, O., Aman, A., Sariyatun, S., Ningsih, T. Z., & Abidin, N. F. (2022). The development of historical thinking assessment to examine students' skills in analyzing the causality of historical events. *European Journal of Educational Research*, *8*(3), 753–761. https://doi.org/https://doi.org/10.12973/eu-jer.11.2.609

[21] Ofianto, O., Erniwati, E., & Fitrisia, A. (2023). Can Primary Source Evidence Skills be Improved by using Virtual Field Trip Learning Resources? *AL-ISHLAH: Jurnal Pendidikan*, *15*(4), 5972–5978. https://doi.org/10.35445/alishlah.v15i4.2620

[22] Pajriah, S., & Suryana, A. (2021). Local Wisdom-Based Character Values on Kawali's Inscription In History Learning. *Paramita: Historical Studies Journal*, *31*(2), 259–269. https://doi.org/10.15294/paramita.v31i2.25752

[23] Pellegrini, M., & Vivanet, G. (2021). Evidence-Based Policies in Education: Initiatives and Challenges in Europe. *ECNU Review of Education*, *4*(1), 25–45. https://doi.org/10.1177/2096531120924670

[24] Pratama, Y., Sariyatun, S., & Joebagio, H. (2017). The development of Means-Ends Analysis and Value Clarification Technique Integration Model to explore the local Wisdom in Historical Learning. *Journal of Education and Learning (EduLearn)*, *11*(2), 179–187. https://doi.org/10.11591/edulearn.v11i2.5752

[25] Priamantono, R. (2020). Implementation of Local Wisdom Values of Piil Pesenggiri as Character Education in

Indonesian History Learning. *VNU Journal of Science: Education Research*, *36*(4), 1–10. https://doi.org/10.25073/2588-1159/vnuer.4366

[26] Putro, H. T., & Setyowati, E. (2022). Development of Application Based on Augmented Reality as A Learning of History and Culture in Architecture Case Study Pathok Negoro Mosques Yogyakarta. *Journal of Artificial Intelligence in Architecture*, *1*(1), 1–9. https://doi.org/10.24002/jarina.v1i1.4835

[27] Reisman, A., Brimsek, E., & Hollywood, C. (2019). Assessment of Historical Analysis and Argumentation (AHAA): A New Measure of Document-Based Historical Thinking. *Cognition and Instruction*, *37*(4), 534–561. https://doi.org/10.1080/07370008.2019.1632861

[28] Salbella, M. W., & Kumalasari, D. (2020). History Learning Based on Local Wisdom of Taluba. *In 2nd International Conference on Social Science and Character Educations (ICoSSCE 2019)*, 107–109. https://doi.org/10.2991/assehr.k.200130.023

[29] Seixas, P., Gibson, L., & Ercikan, K. (2015). A Design Process For Assessing Historical Thinking: The Case of a One-Hour Test. In K. Ercikan & P. Seixas (Eds.), *New Directions in Assessing Historical Thinking* (pp. 124–138). Routled.

[30] Smith, M. D. (2017). Cognitive Validity: Can Multiple-Choice Items Tap Historical Thinking Processes? *American Educational Research Journal*, *54*(6), 1256–1287. https://doi.org/10.3102/0002831217717949

[31] Tally, B., & Goldenberg, L. B. (2005). *Fostering Historical Thinking With Primary Sources*. *5191*, 1–21.

[32] Tirado-Olivares, S., Cózar-Gutiérrez, R., López-Fernández, C., & González-Calero, J. A. (2023). Training future primary teachers in historical thinking through error-based learning and learning analytics. *Humanities and Social Sciences Communications*, *10*(1), 1–11. https://doi.org/10.1057/s41599-023-01537-w

[33] Ufie, A., Matitaputy, J. K., & Kufla, J. (2020). Vean tradition as a local wisdom of customary people and its relevance to maritime history learning. *Journal of Education and Learning (EduLearn)*, *14*(4), 590–598. https://doi.org/10.11591/edulearn.v14i4.16401

[34] Vijayakumar, S., & Ahmad, A. (2023). The Relationship between History Teacher's Competence in Using Digital History Resources and the Inculcation of Historical Thinking Skills. *International Journal of Academic Research in Progressive Education and Development*, *12*(1). https://doi.org/10.6007/ijarped/v12-i1/16393

[35] Wilke, M., Depaepe, F., & Van Nieuwenhuyse, K. (2023). Fostering Secondary Students' Historical Thinking: A Design Study in Flemish History Education. *Journal of Formative Design in Learning*, *7*(1), 61–81. https://doi.org/10.1007/s41686-023-00074-8

[36] Yefterson, R. B., Naldi, H., Lionar, U., & Syafrina, Y. (2020). The Relevance of Local Historical Events in Building National Identities: Identification in the History Learning Curriculum in Indonesia. *International Journal of Progressive Sciences and Technologies (IJPSAT*, *23*(1), 500–504. http://ijpsat.ijsht-journals.org