**Journal of Research and Advances in Mathematics Education**

Volume 11, Issue 2, April 2026, pp. 155-176
DOI: 10.23917/jramathedu.v11i2.15181
p-ISSN: 2503-3697, e-ISSN: 2541-2590

# Developing PISA-like task using climate change contexts to enhance students' statistical literacy

**Duano Sapta Nusantara[1*], Satria Bayu Herlambang[1], Feri Tiona Pasaribu[1], Kgaladi Maphutha[2]**

*[1]Universitas Jambi, Indonesia*
*[2]University of Limpopo, South Africa*

**Citation:** Nusantara, D. S., Herlambang, S. B., Pasaribu, F. T., & Maphutha, K. (2026). Developing PISA-like task using climate change contexts to enhance students' statistical literacy. JRAMathEdu (Journal of Research and Advances in Mathematics Education), 11(2), 155-176. https://doi.org/10.23917/jramathedu.v11i2.15181

**ARTICLE HISTORY:**
Received 29 December 2025
Revised 19 Maret 2026
Accepted 20 April 2026
Published 30 April 2026

**KEYWORDS:**
Climate change
Design research
PISA-like tasks
Statistical literacy
Uncertainty and data

**ABSTRACT**

This study developed PISA-like tasks in the Uncertainty and Data domain using climate change contexts and examined their potential to elicit junior high school students' statistical literacy. The study employed a design research methodology of the development studies type, in which Tessmer's formative evaluation model was applied specifically to the formative evaluation stage, encompassing self-evaluation, expert review, one-to-one, small group, and field test, alongside a preliminary phase and an assessment phase. A total of 49 Grade IX students from a state junior high school in Jambi, Indonesia, participated across the evaluation stages: three in one-to-one, nine in small group, and 37 in the field test. Twenty-two climate change–related tasks were developed. However, this article specifically focuses on one scenario concerning "Jakarta is Singking" due to climate change impacts. Data were collected through expert validation, student practicality questionnaires, written task responses, and interviews, and were analyzed qualitatively and descriptively. The developed tasks were valid in terms of content, construct, and language, and practical based on student practicality scores. The assessment phase revealed a positive potential effect: Problem Understanding reached 76.06%, Data Processing 69.26%, and Data Interpretation 68.69%, all in the moderate category, with interpretation being the least developed due to students' limited prior exposure to context-rich tasks. The climate change context supported engagement with trends, variability, and uncertainty. These findings suggest that future studies should integrate climate-based PISA-like tasks within structured instructional phases to support the written expression of students' statistical reasoning, which in this study emerged more clearly through interviews than written responses.

## INTRODUCTION

In the twenty-first century, statistical literacy has become a fundamental competence as individuals are increasingly required to engage with large volumes of data characterized by uncertainty, variability, and real-world complexity. Statistical literacy is not limited to the ability to perform routine calculations but encompasses the capacity to understand data representations, critically evaluate information, and draw reasoned conclusions for informed decision-making (Büscher, 2022; Chick & Pierce, 2012; Delport, 2023; Gal, 2002; Gould, 2017). This importance is formally recognized in Indonesia's national curriculum; under the Merdeka Curriculum, the Data Analysis and Probability element in Phase D explicitly targets students' ability to validate statements, compare data, determine trends, and make data-based predictions (Kemendikbudristek, 2025),

*Corresponding author: duanosaptanusantara@unja.ac.id

positioning statistical literacy as a central learning outcome for junior high school mathematics. From a theoretical standpoint, statistical literacy involves interconnected processes of problem understanding, data processing, and data interpretation, which enable learners to make sense of quantitative information embedded in everyday and societal contexts (Britton & Anderson, 2020; Utari et al., 2025a). Consequently, strengthening students' statistical literacy is widely recognized as a central goal of contemporary mathematics education (Prahmana & D'Ambrosio, 2020).

One of the most influential international frameworks for assessing these competencies is the Programme for International Student Assessment (PISA), administered by the Organisation for Economic Co-operation and Development (OECD). Evidence from a preliminary study conducted at a junior high school in Jambi revealed comparable patterns, where, of 40 students tested, only 17.5% demonstrated high statistical literacy, while 37.5% were at the moderate level and 45% at the low level; many students relied on This domain is therefore directly aligned with the statistical literacy competencies targeted by the Indonesian curriculum, as both require students to move beyond procedural calculation toward contextual data interpretation and evidence-based reasoning. However, results from successive PISA cycles reveal persistent challenges for Indonesian students. National performance remains below the OECD average, with most students positioned at or below Level 1, while Level 2 represents the minimum expected proficiency for lower secondary students (OECD, 2023). These findings indicate that many students struggle to interpret even basic data displays, identify trends, and justify conclusions based on evidence, suggesting weaknesses in statistical reasoning rather than mere computational skills (Deda & Maifa, 2021; Monteiro & Carvalho, 2023; OECD, 2023; Sihlangu et al., 2025).

Empirical studies further corroborate these findings by highlighting the role of assessment practices in shaping students' statistical literacy (Murdaningsih & Murtiyasa, 2016; Ünal et al., 2023; van Dijke-Droogers et al., 2022; Weiland, 2017). Ramli et al. (2022) reported that classroom assessments in Indonesia are predominantly procedural and routine, offering limited opportunities for students to engage in data interpretation and reasoning within meaningful contexts. Similarly, Sujadi et al. (2023) and (Nusantara, et al., 2025) found that students frequently commit conceptual and interpretative errors when solving PISA problems in the Uncertainty and Data content, particularly when tasks require connecting data to contextual information. Evidence from a preliminary study conducted at a junior high school in Jambi revealed comparable patterns, where, of 40 students tested, only 17.5% demonstrated high statistical literacy, while 37.5% were at the moderate level and 45% at the low level; many students relied on intuitive estimations without adequate reasoning or calculations when interpreting graphs and trends (Kurnia et al., 2023). These conditions underscore the urgent need for assessment instruments that are theoretically aligned with the nature of statistical literacy and capable of eliciting students' data-based reasoning (Callaman & Itaas, 2020; Gonda et al., 2022; Johannssen et al., 2021; Ulia et al., 2023).

In response to these challenges, PISA-like tasks have been increasingly explored as an assessment approach that emphasizes contextual problem solving and statistical literacy. Research by Nusantara et al. (2021) demonstrates that carefully selected contexts can meaningfully support students' engagement with data and interpretation of trends, including the use of COVID-19 transmission maps to reason about change, digital representations to enhance interaction with contextual information (Nusantara et al., 2024a), and analyses of students' strategies when solving problems on change and relationship content (Nusantara et al., 2024b). More recently, the JUMPISA framework provided empirical evidence that systematically designed, context-based PISA tasks can foster deeper statistical reasoning (Tanujaya et al., 2023). In parallel, studies on statistical literacy highlight the importance of coherent task design and learning trajectories that integrate meaningful contexts, showing positive effects on students' statistical literacy development and alignment between statistical concepts and real-world problem situations (Prahmana & D'Ambrosio, 2020; Utari et al., 2025a, 2025b).

Despite these advances, existing studies have largely concentrated on specific contexts such as health-related data, digital environments, or local cultural settings, with relatively limited attention given to globally urgent and data-intensive issues within the Uncertainty and Data domain. In Indonesia, prior PISA-like task development has predominantly emphasized mathematical reasoning

(Islamirta et al., 2022), higher-order thinking (Utari et al., 2024), or thematic contexts such as traditional food (Gustiningsi et al., 2023), historical artifacts (Aini et al., 2022), and the COVID-19 pandemic (Nusantara et al., 2021). Although these studies contributed meaningfully to contextual mathematics education, they have not sufficiently addressed the development of statistical literacy through contexts that inherently require students to interpret uncertainty, variability, trends, and data-driven evidence.

In this regard, climate change is not merely an alternative thematic context, but a highly relevant and authentic domain for fostering statistical literacy within the Uncertainty and Data content. Climate change issues are fundamentally grounded in the interpretation of complex and uncertain data, including long-term temperature trends, carbon emission patterns, probability of extreme weather events, and predictive climate models (IPCC, 2023). Such characteristics closely align with the core competencies of statistical literacy, particularly the ability to interpret data, evaluate variability, make evidence-based conclusions, and reason under conditions of uncertainty. However, climate change remains underutilized in PISA-like task development, especially within the Indonesian context. The few studies that have incorporated climate-related contexts, such as Utari et al. (2024), who used global warming data to promote higher-order thinking, and Kurniadi et al., (2026), who demonstrated that climate change data can engage students in critical environmental analysis, focused primarily on mathematical thinking rather than explicitly targeting statistical literacy indicators in the Uncertainty and Data domain.

Therefore, a significant gap remains in the development of PISA-like tasks that integrate climate change as an authentic context for cultivating statistical literacy. Integrating climate-related data into assessment tasks provides meaningful opportunities for students to engage in statistical reasoning while simultaneously addressing global sustainability challenges aligned with the Sustainable Development Goals, particularly Goal 4 (Quality Education) and Goal 13 (Climate Action) (Kurniadi et al., 2026; OECD, 2016). Such integration not only strengthens students' statistical literacy competencies but also enhances their awareness of evidence-based decision making in responding to contemporary global issues.

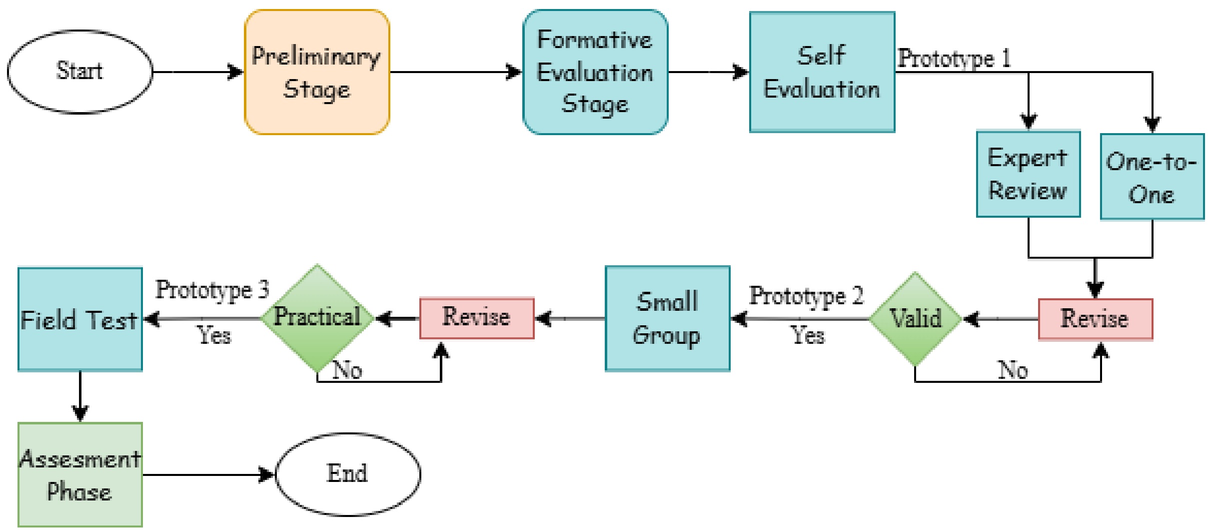
Accordingly, this study aims to develop PISA-like tasks in the Uncertainty and Data domain using climate change contexts through Tessmer's formative evaluation model. This approach is employed to ensure that the developed tasks are valid in terms of content, construct, and language, practical for classroom implementation, and possess potential effects in eliciting junior high school students' statistical literacy. By extending PISA-oriented task development to a globally relevant and data-intensive climate change context, this study addresses a notable gap within the Uncertainty and Data domain. The study contributes theoretically by reinforcing the conceptual integration between statistical literacy, context-based assessment, and sustainability-oriented mathematics education, and contributes practically by producing a validated and empirically examined PISA-like assessment task that is feasible for classroom use and capable of eliciting students' statistical literacy in authentic data contexts.

## METHODS

This section outlines the research methods used in this study, including the research design, participants, implementation procedures, data collection techniques, and data analysis.

### Research design

This study employed a design research methodology of the development studies (Bakker, 2018; Plomp & Nieveen, 2013). The developmental process followed Tessmer's formative evaluation model (Tessmer, 1993), consisting of a preliminary phase, formative evaluation (self-evaluation, expert review, one-to-one, small group, and field test), and an assessment phase (Zulkardi, 2002). This approach was selected to produce PISA-like task in the Uncertainty and Data domain that are valid, practical, and potentially effective for enhancing statistical literacy.




(Zulkardi, 2002)

**Figure 1.** Procedure developing PISA-like task for enhancing statistical literacy

## Research participants

The participants were Grade IX students from a state junior high school in Jambi, Indonesia. Student categorization into low, medium, and high achievement levels was determined based on their mathematics daily test scores, as these scores reflect students' recent mastery of statistical and data-related content. Three students representing each achievement category participated in the one-to-one evaluation stage, while nine students took part in the small-group evaluation. There were 37 students who involved in the field test stage to examine the practicality and potential effects of the developed tasks. In addition, one mathematics teacher provided curricular and classroom-based judgments, while three mathematics education researchers specializing in PISA item development and statistical literacy served as expert validators to evaluate the content validity and construct quality of the tasks.

## Research procedure

Figure 1 presents the overall research procedure adapted from Tessmer (1993) and Zulkardi (2002) formative evaluation model. The diagram illustrates the sequential flow of activities across the preliminary phase, formative evaluation, and assessment phase. As shown in Figure 1, the procedure consisted of three phases. The preliminary phase involved document analysis of the PISA 2022 framework, the Indonesian Merdeka Curriculum (Phase D), relevant PISA items, and prior studies to construct Draft 1 containing item specifications, scoring rubrics, and five climate change contextual themes global warming, deforestation, natural disaster frequency, air temperature anomalies, and Jakarta land subsidence selected for their data richness and alignment with students' lived experiences and the SDGs (Kurniadi et al., 2026). The formative evaluation stage refined the tasks through iterative cycles: self-evaluation produced Prototype 1; expert review generated Prototype 2; one-to-one trials improved readability; small-group trials assessed practicality and led to Prototype 3; and a full-class field test produced the Final Prototype. The assessment phase examined the potential effectiveness of the final tasks by analyzing students' reasoning and the appearance of statistical literacy indicators.

## Data collection

Data were collected through six techniques applied across the development stages. First, document analysis of the PISA 2022 framework, Merdeka Curriculum standards, and prior studies was conducted during the preliminary phase to inform the construction of the initial task draft. Second, walkthroughs were used during the expert review stage, in which validators provided written feedback and recommendations on content, construct, and language quality. Third, validity

**Table 1**
Criteria for validity, practicality, and potential effect analysis

| Aspect | Analysis Procedure | Percentage Range | Qualification |
|---|---|---|---|
| Validity | Validity questionnaires were scored using a five-point Likert scale and converted into percentages of the maximum possible score. | $75 < score \leq 100$<br>$50 < score \leq 75$<br>$25 < score \leq 50$<br>$0 < score \leq 25$ | Very Valid<br>Valid<br>Invalid<br>Very Invalid |
| Praticality | Practicality questionnaires were scored using a five-point Likert scale and converted into percentages. | $61 < score \leq 81$<br>$41 < score \leq 61$<br>$61 < score \leq 81$<br>$21 < score \leq 61$<br>$0 < score \leq 21$ | Practical<br>Fairly Practical<br>Practical<br>Less Practical<br>Not Practical |
| Potential Effect | Students' written responses were analyzed using a partial-credit scoring rubric ranging from 0–2 for each item. A score of 2 indicated a complete and accurate response, 1 indicated a partially correct response, and 0 indicated an incorrect or no response. Scores were aggregated into a final score percentage. | $80 < score \leq 100$<br>$60 < score \leq 80$<br>$0 < score \leq 60$ | High<br>Moderate<br>Low |

questionnaires were administered to the four validators during the expert review stage to obtain quantitative ratings across nine indicators covering content, construct, and language. Fourth, student practicality questionnaires were administered during the small-group and field test stages to assess usability, readability, and feasibility. Fifth, written task responses from 37 field-test students were collected to examine the potential effects of the final set of ten PISA-like tasks on students' statistical literacy. Sixth, semi-structured interviews were conducted with selected students at each evaluation stage at one-to-one to probe comprehension and language clarity, at small group to assess usability, and at field test to explore how the climate change context supported students' data reasoning.

## Data analysis

Data analysis combined qualitative and quantitative techniques at different stages. Qualitatively, walkthrough and interview data were analyzed descriptively to identify issues of content, construct, and linguistic clarity, and observational notes capturing student behaviors such as expressions of confusion, problem-solving strategies, and engagement with data representations were analyzed at each stage to inform task revision.

The quantitative data analysis procedures included the evaluation of content validity, practicality, and the potential effect of the developed PISA-like tasks. The scoring procedures, assessment criteria, and corresponding references used in this study are summarized in Table 1. As shown in Table 1, the validity and practicality analyses were conducted using questionnaire data scored on a five-point Likert scale and converted into percentages of the maximum possible score (Rodriguez-Martinez et al., 2023; Utari et al., 2024). Meanwhile, the potential effect analysis was based on students' written responses using a partial-credit scoring rubric adapted from OECD (2019). The resulting scores were categorized into high, moderate, and low levels to determine the extent to which the developed tasks supported students' statistical literacy performance.

To further examine students' statistical literacy performance, students' responses were analyzed based on three statistical literacy indicators adapted from Schield (2011) and Gal (2024). The indicators and their operational explanations are presented in Table 2. The indicators presented in Table 2 served as analytical categories for identifying patterns of students' statistical reasoning when solving the climate change-based PISA-like tasks. Through these indicators, the analysis focused on how students understood the problem context, processed the provided data, and interpreted the results to produce evidence-based conclusions.

**Table 2**
Statistical literacy indicator used in the analysis

| Indicator | Code | Descriptor |
|---|---|---|
| Problem Understanding | PU | Students correctly identified relevant information, recognized variables, and understood question demands from graphs, tables, or other visual representations. |
| Data Processing | DP | Students selected and applied appropriate strategies, such as calculating differences, identifying rates of change, comparing data, or making predictions based on available data. |
| Data Interpretation | DI | Students drew logical and evidence-based conclusions related to the climate change context and justified their reasoning appropriately. |

**Table 3**
The result of the independent curriculum analysis

| Element | Learning Outcomes | Relation to Statistical Literacy |
|---|---|---|
| Data Analysis and Probability | By the end of Phase D, Students are expected to be able to validate statements, compare data, calculate changes, determine trends, and make simple mathematical predictions. | These learning outcomes align with core statistical literacy competencies in the *Uncertainty and Data* domain, particularly students' ability to interpret and critically evaluate quantitative information, identify patterns and variability in data, analyze trends across datasets, and formulate evidence-based conclusions and predictions within authentic contextual situations. |

## FINDINGS

This section presents the development process and results of the PISA-like assessment tasks designed for the Uncertainty and Data domain to enhance junior high school students' statistical literacy.

### Preliminary phase

The preliminary phase aimed to establish a coherent foundation for task development by aligning student characteristics, curriculum demands, PISA framework requirements, and the selected climate-change contexts. This phase consisted of preparatory analyses followed by the construction of an initial task draft.

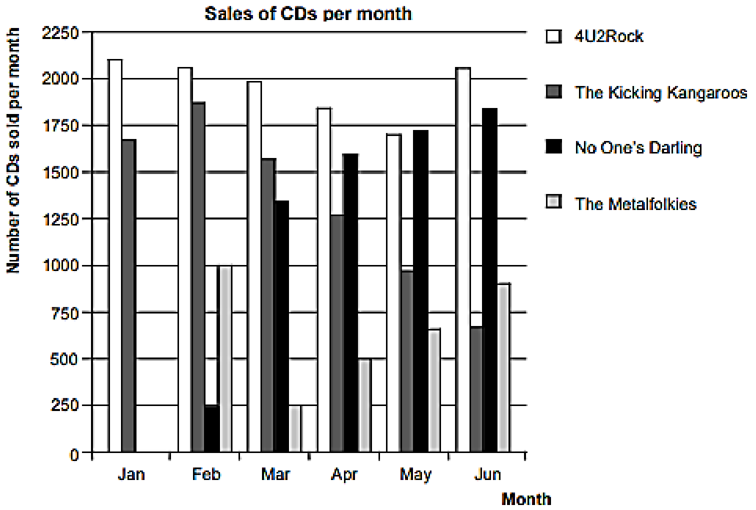
The process began with a student analysis conducted in collaboration with mathematics teachers at Secondary School 6 Kota Jambi. The analysis focused on students' academic diversity and learning readiness to ensure that participants selected for each evaluation stage appropriately represented varying ability levels. Based on this analysis, three Grade IX C students were selected for the one-to-one stage, nine Grade IX B students participated in the small-group evaluation, and one intact Grade IX A class was designated for the field test.

In terms of curriculum, the secondary school implements the *Kurikulum Merdeka* (Independent Curriculum–IC), which is consistent with the curriculum framework adopted by secondary schools across Jambi Province, including schools participating in the development of PISA-like tasks within the *Uncertainty and Data* domain. The IC places substantial emphasis on data literacy and evidence-based reasoning, particularly through competencies requiring students to analyze, evaluate, and interpret quantitative information in contextual situations. These competencies conceptually align with the construct of statistical literacy, which encompasses the ability to interpret data representations, identify patterns and variability, evaluate quantitative claims, and formulate evidence-based conclusions in situations involving uncertainty.

Accordingly, the learning objectives and task indicators developed in this study were aligned with the learning outcomes specified at the end of Phase D of the IC, particularly within the *Data Analysis and Probability* element, as presented in Table 3. The Data Analysis and Probability learning outcomes provide a strong conceptual and curricular basis for the development of statistical literacy-oriented PISA-like tasks. The emphasis on interpreting data, identifying trends, and generating predictions is particularly relevant to the Uncertainty and Data domain, as these competencies

CHARTS

In January, the new CDs of the bands *4U2Rock* and *The Kicking Kangaroos* were released. In February, the CDs of the bands *No One's Darling* and *The Metalfolkies* followed. The following graph shows the sales of the bands' CDs from January to June.




**Question 5: CHARTS** PM918Q05

The manager of *The Kicking Kangaroos* is worried because the number of their CDs that sold decreased from February to June.

What is the estimate of their sales volume for July if the same negative trend continues?

A. 70 CDs
B. 370 CDs
C. 670 CDs
D. 1340 CDs

**Figure 2.** Inspiration for the developed PISA-like task using climate change context

require students to engage in statistical reasoning beyond procedural computation. Consequently, the curriculum supports the integration of authentic and data-intensive contexts, such as climate change, to facilitate students' capacity to critically interpret information and make informed decisions based on empirical evidence.

To ensure alignment with international assessment standards, the PISA framework was analyzed with particular attention to content, context, and mathematical processes. The content domain emphasized Uncertainty and Data, while the contexts were drawn from scientific and social issues related to climate change. In terms of mathematical processes, the tasks were designed to engage students in the formulate–employ–interpret (FEI) cycle as defined in the PISA framework (OECD, 2023).

As part of this analysis, several released PISA items were examined to understand their structure, use of data representations, and cognitive demands. One example is the *Charts* task from PISA 2012, which presents a bar chart showing monthly CD sales of several music bands and asks students to interpret trends across time (see Figure 2). This task requires students to identify relevant variables, compare values across categories, and reason about changes over successive periods rather than simply performing routine calculations.

Figure 2 The *Charts* task was not reused in this study; instead, it served as a conceptual reference for designing PISA-like task that emphasize trend analysis and interpretation of data over time. Drawing on this structure, the developed tasks in this study were designed to require students to analyze historical data and make predictions about future conditions. For example, students were asked to estimate future temperature values based on trends observed in climate data from previous years.

This design choice intentionally moved beyond routine procedural tasks and encouraged multi-step reasoning and contextual interpretation. Although a formal psychometric level classification was not conducted at this stage, the tasks were constructed to reflect the cognitive characteristics of PISA items that involve reasoning with trends and making data-based predictions. This consideration informed the intended level of task complexity without imposing rigid level claims.

The final in the preliminary phase was the construction of the initial task draft. A test specification table (blueprint) was developed to map curriculum learning outcomes, PISA framework components, statistical literacy indicators, and climate-change contexts. Based on this blueprint, a set of open-ended, constructed-response tasks was written, accompanied by a scoring rubric designed to support consistent and objective assessment of students' responses. The output of this phase was the initial version of the developed instrument, referred to as Prototype I, which served as the basis for subsequent refinement during the formative evaluation phase.

## Formative evaluation stage

The formative evaluation stage was conducted through a formative evaluation cycle aimed at refining and validating the developed PISA-like tasks. This stage consisted of several iterative stages, namely self-evaluation, expert review, one-to-one evaluation, small-group evaluation, and field testing. Each stage played a specific role in improving the quality of the tasks in terms of content validity, clarity, contextual relevance, and their ability to elicit students' statistical literacy within the Uncertainty and Data domain.

### *Self-evaluation*

At the self-evaluation stage, the researcher conducted an internal review of Prototype I by referring to the PISA framework, particularly the Uncertainty and Data domain and the FEI mathematical process (OECD, 2023). This stage aimed to ensure that the task was conceptually aligned with statistical literacy indicators and emphasized reasoning with data and uncertainty rather than routine procedural calculations.

The task was developed using an informational poster presenting projected data on the percentage of Jakarta's area potentially inundated by seawater due to continued land subsidence. The "Jakarta Tenggelam" context was selected as an authentic climate change is arelated issue involving uncertainty, trend analysis, and long-term data interpretation. Students were asked to estimate the year in which 50% of Jakarta's area might be submerged based on the observed data trend and to justify their prediction using evidence from the poster.

During self-evaluation, the researcher examined whether the task required students to identify relevant quantitative information, recognize trends in the projected data, and extrapolate these trends to make a reasoned prediction. Particular attention was given to engaging the *interpret* phase of the PISA mathematical process, with the task intentionally designed to avoid direct formula application and instead emphasize data interpretation and justification under uncertainty.

Based on this evaluation, several revisions were made, including refining the wording for clarity, strengthening the conceptual link between land subsidence and climate change, and ensuring that the prompt explicitly directed students to base their predictions on data trends rather than subjective opinion. Visual elements were also reviewed to support clearer comparison across years.

Figure 3 illustrates the revised Prototype I following the self-evaluation stage, highlighting the improved clarity of the problem statement and the visual presentation of projected inundation data that supports students' trend analysis and prediction. Figure 3 presents Prototype I of the developed PISA-like task, which uses the "Jakarta is sinking?!?" context as a representative example of the climate change theme. This task presents an informational poster containing projected data on the percentage of Jakarta's area potentially inundated by seawater due to continued land subsidence, a phenomenon driven by a combination of excessive groundwater extraction, increased load from tall buildings, and rising sea levels resulting from climate change. Students are asked to estimate the year in which 50% of Jakarta's area will be flooded and to justify their prediction mathematically based on the observed data trend. This context was selected because it inherently involves uncertainty,

**Jakarta is sinking?!?**

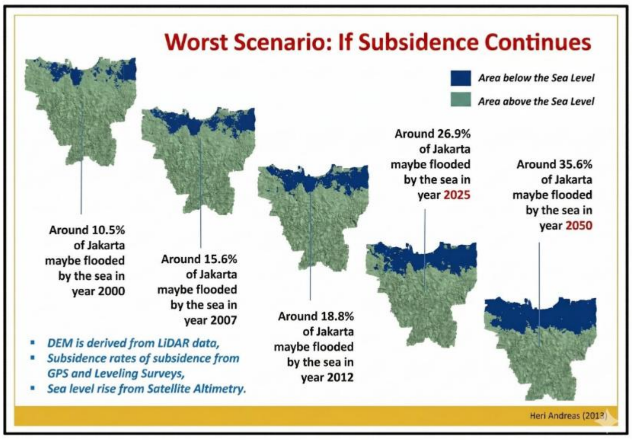
Did you know that Jakarta is one of the cities with the fastest land subsidence in the world?

https://www.detik.com/edu/detikpedia/d-7246014/jakarta-jadi-kota-yang-paling-cepat-tenggelam-di-dunia-mengapa

This phenomenon is caused by a combination of various factors, including:

- Excessive groundwater extraction,
- Increased load from tall buildings,
- Rising sea levels due to climate change.

As a result, parts of Jakarta are already below sea level and are predicted to expand in the future. The data below shows the projected percentage of Jakarta that may be flooded by seawater if land subsidence continues.




**Question:**

Based on the data above, it is estimated that 50% of Jakarta will be flooded by seawater in the future. In your opinion, in what year might this condition occur? Explain your prediction using the available data trends.

**Figure 3.** Prototype I of PISA-like task using climate change context

trend analysis, and long-term data interpretation, which are central characteristics of the Uncertainty and Data domain, while also representing a locally and globally relevant climate issue that is meaningful to Indonesian students.

### *Expert review*

The expert review stage was conducted to obtain systematic feedback and recommendations on the quality of Prototype I, with particular emphasis on content validity, construct validity, and language clarity. This stage aimed to ensure that the developed PISA-like task met essential quality standards before proceeding to subsequent formative evaluation phases. A summary of the experts' quantitative and qualitative evaluations is presented in Table 2 and Table 3.

Four experts participated in this stage, consisting of three university lecturers in mathematics education and one experienced junior high school mathematics teacher. The lecturers have extensive experience in mathematics education research and PISA-oriented task development, including the design and evaluation of PISA-like assessment items, while the teacher has more than 15 years of classroom teaching experience and has been involved in school-based assessment practices. This combination of academic and practical expertise ensured a comprehensive evaluation of the task from both theoretical and classroom perspectives.

The quantitative results of the expert validation are summarized in Table 4. The validation was conducted using a five-point Likert scale (1 = strongly disagree to 5 = strongly agree) across nine indicators covering content, construct, and language aspects. The validation results indicate that Prototype I achieved an average validity percentage of 83.9%, which falls within the *very valid* category (Utari et al., 2024). This result demonstrates that the developed tasks were considered

**Table 4**

Validation scores by expert review on PISA-like tasks with a climate change context

| Validation Criteria | V1 (Obtained/Max Score) | V2 (Obtained/Max Score) | V3 (Obtained/Max Score) | V4 (Obtained/Max Score) | Average (%) | Categor |
|---|---|---|---|---|---|---|
| Content relevance and alignment with statistical literacy competencies | 35/45 (77,8%) | 40/45 (88,9%) | 32/45 (71%) | 44/45 (97.8%) | 83,9 | Very Valid |

Note: *Validation scores were obtained from expert judgments using a five-point Likert scale and converted into percentages of the maximum possible score. The validity category was determined based on the criteria presented in Table 1.*

**Table 5.**

Validator feedback on PISA-like tasks with a climate change context

| Validator (V) | Comments and Suggestions |
|---|---|
| V1 | replace the term *increase* with a more appropriate expression, as an increase inherently implies upward movement and cannot be described as slowing, even if the change is small. |
| V2 | The item is considered usable with minor revisions. |
| V3 | The item can be implemented with slight revisions, including the addition of clear titles or captions for images or tables (e.g., *"Figure 1. ..."*). Furthermore, the term *"trends"* should be paraphrased. |
| V4 | It is recommended to avoid using bold formatting in the question statements to ensure that students. |

appropriate by the validators across content, construct, and language dimensions. From the content perspective, the experts agreed that the tasks were consistent with the PISA 2022 *Uncertainty and Data* framework and successfully integrated statistical literacy indicators, particularly problem understanding, data processing, and data interpretation, through authentic climate change situation. The validators highlighted that the use of temperature trends, emission data, and environmental statistics provided meaningful opportunities for students to engage in evidence-based reasoning and interpretation of uncertainty-related information. In terms of contruct validity, the validators considered the organization of the task components, including the stimulus presentation, question sequencing, and scoring rubric, to be coherent and capable of eliciting the intended statistical reasoning processes. The contextual stimuli were judged to sufficiently support students in identifying relevant information, analyzing quantitative relationships, and constructing data-based conclusions. Nevertheless, several revisions were recommended, particularly regarding the refinement of graphical representations and the simplification of several question prompts to improve clarity and reduce potential misinterpretation. From the language perspective, the validators agreed that the language used in the tasks was generally appropriate for junior secondary school students and aligned with students' cognitive characteristics. Minor revisions were suggested to improve sentence effectiveness, consistency of terminology, and readability of instructions. These revisions were subsequently incorporated into Prototype II, resulting in tasks that were not only theoretically aligned with the statistical literacy framework but also practically suitable for classroom implementation.

**Table 6**
Student comments and responses during the one-to-one evaluation

| Students (S) | Comment and Suggestion |
|---|---|
| S1 | No comments or suggestions were provided |
| S2 | students experienced difficulty understanding the term *"trend"* |
| S3 | students experienced difficulty understanding the term *"trend" and the text before the image is very long* |

In addition to numerical ratings, the experts provided written comments and suggestions that offered deeper insight into specific aspects requiring refinement. These qualitative findings are summarized in Table 5.

Regarding language clarity, experts recommended improving sentence formulation to enhance student comprehension, including correcting punctuation usage, such as limiting the question mark in the item title *"Jakarta is singking ?!?"* and paraphrasing terms like *"trend,"* which were considered potentially difficult for students at the junior secondary level. From a construct perspective, experts suggested adding clear titles or captions to images and tables (e.g., *"Figure 1. ..."*) to support accurate interpretation of visual representations. Furthermore, they advised avoiding bold formatting in question statements to ensure that students rely on data interpretation and visual literacy rather than textual emphasis.

Overall, although Prototype I was judged to be very valid, the experts' suggestions provided valuable guidance for refinement. These recommendations were used as the basis for revising Prototype I to improve clarity, readability, and alignment with students' mathematical literacy skills, particularly in interpreting data and representations within a climate change context.

### *One-to-one*

The one-to-one evaluation was conducted to identify students' difficulties in understanding the task, particularly in relation to language use and data interpretation. Overall, students were able to recognize the climate change issue embedded in the task and attempted to make predictions based on the data provided. This finding suggests that the task was generally accessible and capable of eliciting students' statistical reasoning within the *Uncertainty and Data* domain.

Table 6 presents students' comments and responses during the one-to-one evaluation. While one student reported no difficulties, two students indicated problems in understanding the term *"trend."* These responses highlight that, although students could interpret the visual data and observe numerical changes, specific terminology posed challenges that affected their reasoning process.

Further qualitative analysis revealed that the main difficulty was linguistic rather than mathematical. In particular, unfamiliarity with the term *"trend"* limited students' ability to justify their predictions mathematically. Students also noted that clearer visual labels and more explicit question wording would better support their reasoning. This issue is illustrated in the following interaction between the researcher I and a student (S):

*R : What do you think about the question? Is there any part that is difficult to understand?*
*S : I understand the picture and the data, but I'm confused by the word "trend." I'm not sure what I'm supposed to look for.*
*R : So how did you decide your answer?*
*S : I just guessed based on the numbers going up, but I wasn't sure if that was what the question wanted.*

This interaction shows that the students' difficulty did not stem from interpreting numerical data, but from understanding specific terminology that guides statistical reasoning; therefore, based on the results of the expert review and one-to-one evaluation, several revisions were implemented, including simplifying the introductory paragraph to help students focus on the data, adding clear labels to visual representations, revising visual information to ensure accurate geographic naming, and refining the question wording to explicitly guide mathematical reasoning, namely: *"Estimate the*

**Jakarta Will Sink**

According to detik.com, Jakarta is one of the cities with the fastest land subsidence in the world. This phenomenon is caused by a combination of factors, including excessive groundwater extraction, increased load from tall buildings, and sea level rise due to climate change.

As a result, parts of Jakarta are already below sea level and are predicted to expand in the future. The data below shows the projected percentage of Jakarta that may be flooded by seawater if land subsidence continues.

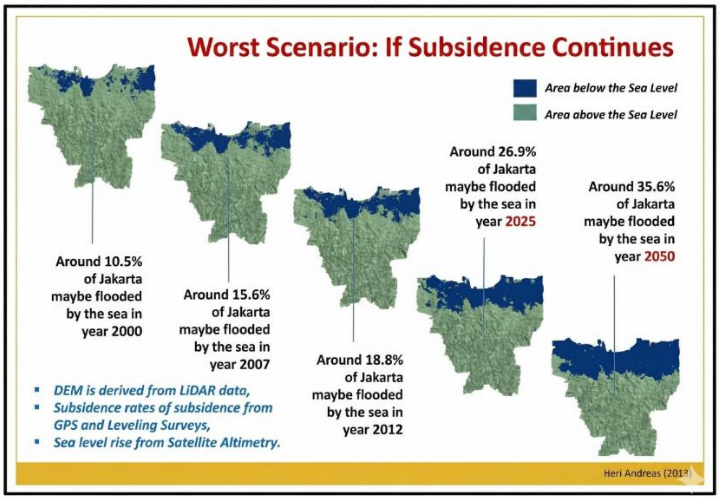



Figure 4. Prediction of Jakarta Area from Year to Year

Source:https://www.detik.com/edu/detikpedia/d-7246014/jakarta-jadi-kota-yang-paling-cepat-tenggelam-di-dunia-mengapa

**Question:**

Estimate the year in which 50% of Jakarta's area will be inundated by seawater and explain your reasoning mathematically!

**Figure 4.** Prototype II of the developed PISA-like task using a climate change context

*year in which 50% of Jakarta's area will be inundated by seawater and explain your reasoning mathematically."* These findings indicate that despite the linguistic challenges, students demonstrated emerging statistical reasoning: they were able to read data representations, identify relevant patterns, and attempt data-based predictions, as evidenced by their written responses showing recognition of trends across time (see Figure 4), which confirms that the task is capable of eliciting statistical reasoning even at the readability refinement stage. These revisions were made to improve language clarity and ensure that students' predictions were supported by data-based reasoning. The revised task was finalized as Prototype II.

### *Small group*

Following the revisions made to Prototype I based on feedback from the expert review and one-to-one evaluation stages, Prototype II was developed and subsequently tested in the small group stage. This stage aimed to examine the practicality of the task, particularly its clarity, usability, and feasibility when implemented with multiple students in a classroom-like setting.

Prototype II was administered to nine junior high school students with varied levels of mathematical ability. Students were given two class period (equivalent to 70 minutes) to complete the task. During the implementation, the researcher closely observed students' interactions with the task, noted questions raised by students, and conducted brief interviews to collect their comments, suggestions, and perceived difficulties while solving the problem.

Overall, most students were able to understand the task instructions and engage with the climate change context appropriately. Students demonstrated the ability to read the provided data, recognize patterns, and attempt to justify their predictions. These observations indicate that the task

**Table 7**
The results of practicality of the developed PISA-like task

| No | Student | Score Obtained | Maximum Score | Percentage (%) |
|---|---|---|---|---|
| 1 | AFS | 59 | 60 | 98,3 |
| 2 | ANF | 43 | 60 | 71,6 |
| 3 | ADLDH | 52 | 60 | 86,6 |
| 4 | AMKAB | 43 | 60 | 71,6 |
| 5 | NPR | 49 | 60 | 81,6 |
| 6 | HA | 54 | 60 | 90 |
| 7 | RZBA | 47 | 60 | 78,3 |
| 8 | KRS | 34 | 60 | 56,7 |
| 9 | SAJ | 45 | 60 | 75 |
| | | Average | | 78,8 |
| | | Category | | Practical |

was generally practical and could be implemented within the allocated instructional time. However, several students encountered difficulties in producing accurate solutions, which were mainly related to limited prior knowledge and unfamiliarity with PISA-like problem formats rather than issues with task clarity or language.

Based on these findings, the researcher concluded that no further revisions were necessary at this stage. The difficulties observed were considered reasonable given students' prior learning experiences and did not indicate weaknesses in the task design. Therefore, Prototype II was retained in its current form for the subsequent evaluation stage.

Table 7 presents the results of the practicality questionnaire administered during the small group evaluation, showing that the developed task achieved an average practicality score of 78.8%, which falls into the "Practical" category according to the established criteria (Rodriguez-Martinez et al., 2023). This finding indicates that the PISA-like task was well received by students in terms of content presentation, clarity of language, and functional usability, confirming that Prototype II can be effectively implemented in classroom settings and is suitable for use in the subsequent evaluation phase.

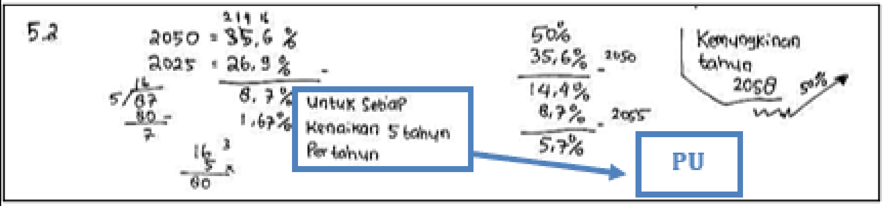
### Field test

The field test stage was conducted to examine the potential effect of the developed PISA-like tasks on students' statistical literacy in an authentic classroom setting. At this stage, the focus was not on measuring learning gains but on identifying whether the tasks could elicit students' statistical literacy as intended within the Uncertainty and Data domain.

The field test was implemented in Grade IX A of Secondary School 6 Kota Jambi. Students worked individually on the Final Prototype, which had undergone expert review, one-to-one, and small group evaluations. Students' written responses were analyzed qualitatively using three indicators of statistical literacy: Problem Understanding (PU), Data Processing (DP), and Data Interpretation (DI). To illustrate the emergence of these indicators, responses from two representative students, coded RAM and ASFZ, are presented. These students were selected on the basis of the clarity and contrasting depth of statistical reasoning demonstrated in their written responses rather than by achievement category, as the analytical focus of this stage was to illustrate how the tasks elicit statistical literacy indicators across varying reasoning patterns.

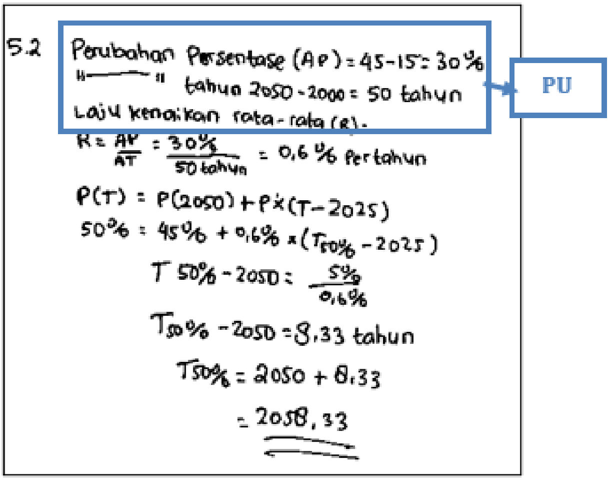
### Problem understanding (PU)

The Problem Understanding (PU) indicator refers to students' ability to comprehend the problem situation, identify relevant information, and recognize what is being asked. Figure 5 presents RAM's response to the *Jakarta is singking* task. The student correctly identified that the problem required predicting the year in which 50% of Jakarta's area would be inundated. RAM recognized that the prediction should be based on the trend shown in the data, indicating an




**English Version**
For every 5-year increase
per year
Possibly in the year 2058

**Figure 5.** PU indicators in RAM students' answers to the question "*Jakarta is Sinking*"




**English Version**
Percentage Change (ΔP): 45 − 15 = 30%
Time Interval (ΔT): 2 050 − 2000 = 50 *years*
Average Growth Rate (R): $R = \frac{\Delta P}{\Delta T} = \frac{30\%}{50\ years} = 0.6\%\ per\ year$
Projection Equation P(T): $P(T) = P(2050) + R \times (T - 2050)$
Finding the Year when Percentage reaches 50%: $50\% = 45\% + 0.6\% \times (T_{50\%} - 2050)$
$T_{50\%} - 2050 = \frac{5\%}{0.6\%}$
$T_{50\%} - 2050 \approx 8.33\ years$
$T_{50\%} = 2050 + 8.33$
$T_{50\%} = 2058.33$

**Figure 6.** PU indicators in ASFZ's answer to the "Jakarta is Sinking" task

understanding that the task involved analyzing changes over time rather than describing isolated values. This understanding was confirmed through a short interview:

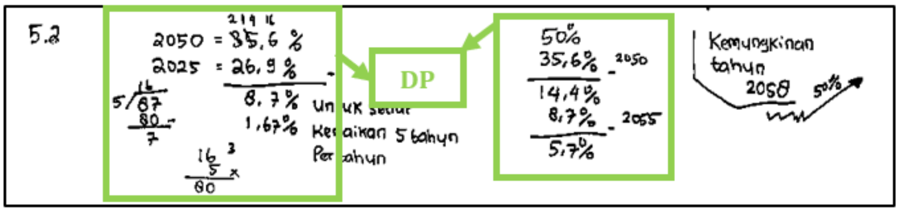
*R* : *What does this question ask you to find?*
*RAM* : *It asks when 50% of Jakarta will be flooded, based on the data.*

This response shows that RAM clearly understood both the context and the objective of the task. The PU indicator emerged clearly, indicating that the task successfully guided students to understand the problem context and the nature of the required prediction.

Figure 6 presents ASFZ's response to the same task. ASFZ demonstrated an understanding of the concept being asked, namely calculating the average rate of increase and applying a linear approach to predict the year in which the percentage of inundated area would reach 50%. Although the numerical values used were not fully consistent with the data in the figure, the reasoning steps demonstrated were already appropriate. This was confirmed through interview:

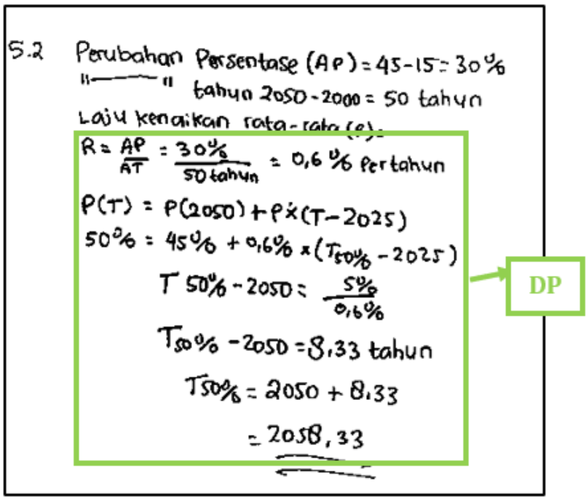
*R* : *What steps did you take to predict the year?*
*ASFZ* : *I took the data, calculated the average per year, then put it into the formula.*

This response shows that ASFZ understood the basic modeling steps required for making a prediction. The PU indicator emerged very clearly on ASFZ, indicating a strong grasp of both what the task demanded and how data should be used to address it.




**English Version**
For every 5-year increase per year
Possibly in the year 2058

**Figure 7.** DP indicators in RAM students' answers to the question "Jakarta is Sinking"




**English Version**
Percentage Change ($\Delta P$): $45 - 15 = 30\%$
Time Interval ($\Delta T$): $2\ 050 - 2000 = 50\ years$
Average Growth Rate (R): $R = \frac{\Delta P}{\Delta T} = \frac{30\%}{50\ years} = 0.6\%\ per\ year$
Projection Equation P(T): $P(T) = P(2050) + R \times (T - 2050)$
Finding the Year when Percentage reaches 50%: $50\% = 45\% + 0.6\% \times (T_{50\%} - 2050)$
$T_{50\%} - 2050 = \frac{5\%}{0.6\%}$
$T_{50\%} - 2050 \approx 8.33\ years$
$T_{50\%} = 2050 + 8.33$
$T_{50\%} = 2058.33$

**Figure 8.** DP indicators in ASFZ's answer to the "Jakarta is Sinking" task

## Data Processing (DP)

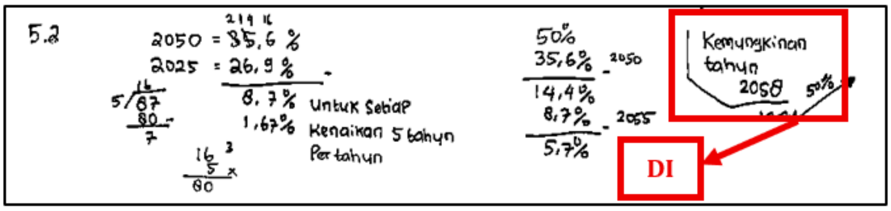
The Data Processing (DP) indicator focuses on students' ability to process numerical information, select strategies, and perform calculations or estimations to support their reasoning. Figure 7 shows RAM's written work when attempting to determine the predicted year. The student identified an increasing pattern in the data and attempted to extend this pattern to reach the 50% threshold. However, RAM relied on informal estimation rather than a systematic calculation of the rate of increase, resulting in an inaccurate prediction.This reasoning was clarified during the interview:

*R* : *How did you decide on that year?*
*RAM* : *I estimated it from how the numbers keep going up each year.*

Although the calculation was not fully accurate, RAM engaged in data processing by using the available data to make an estimate rather than guessing randomly. The DP indicator appeared in the student's response, showing engagement with data-based reasoning, even though the applied strategy was still intuitive and not fully structured.

Figure 8 shows ASFZ's written work for the same task. ASFZ demonstrated a substantially more systematic approach, calculating the total increase of approximately 30% over a 50-year span, dividing it to obtain an annual rate of increase, and then applying linear extrapolation to predict when the inundated area would reach 50%. The error that occurred was only in the selection of the initial value, not in the data processing procedure itself. This was confirmed through interview:

*R* : *How did you determine the year when it would reach 50%?*
*ASFZ* : *I calculated the total increase first, divided it by the number of years, then used that to predict when it would reach 50%.*



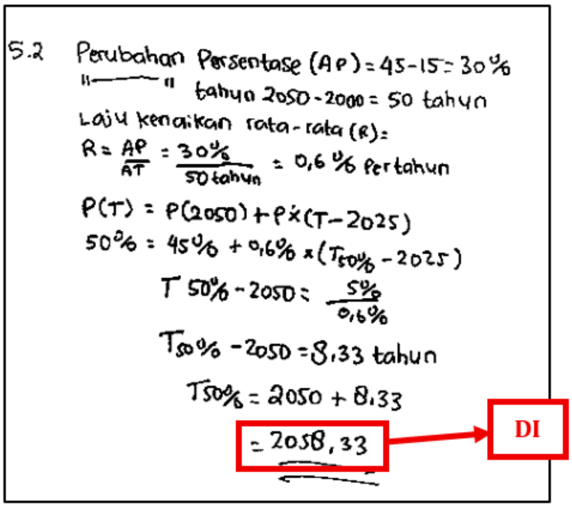

**English Version**

For every 5-year increase

per year

Possibly in the year 2058

**Figure 9.** DI indicators in RAM students' answers to the question "Jakarta is sinking"




**English Version**

Percentage Change ($\Delta P$): $45 - 15 = 30\%$

Time Interval ($\Delta T$): $2\ 050 - 2000 = 50\ years$

Average Growth Rate (R): $R = \frac{\Delta P}{\Delta T} = \frac{30\%}{50\ years} = 0.6\%\ per\ year$

Projection Equation P(T): $P(T) = P(2050) + R \times (T - 2050)$

Finding the Year when Percentage reaches 50%: $50\% = 45\% + 0.6\% \times (T_{50\%} - 2050)$

$T_{50\%} - 2050 = \frac{5\%}{0.6\%}$

$T_{50\%} - 2050 \approx 8.33\ years$

$T_{50\%} = 2050 + 8.33$

$T_{50\%} = 2058.33$

**Figure 10.** DI indicators in ASFZ's answer to the "Jakarta is Sinking" task

Based on both the written response and the interview, ASFZ carried out the data processing in a sequential and logical manner. The DP indicator emerged strongly, reflecting a structured mathematical approach to prediction under uncertainty.

## Data interpretation (DI)

The Data Interpretation (DI) indicator evaluates students' ability to draw conclusions, justify predictions, and relate results to the real-world context. Figure 9 presents RAM's conclusion. The student provided a predicted year and justified it by referring to the observed increase in inundated area. While the explanation was brief and not mathematically rigorous, RAM acknowledged that the prediction was an estimate based on data trends rather than an exact value. This was reflected in the interview response:

| | | |
|---|---|---|
| *R* | : | *Why do you think your answer makes sense?* |
| *RAM* | : | *Because the data keeps increasing, so if it continues, it will reach 50%.* |

This explanation demonstrates an emerging ability to interpret data and communicate conclusions within a contextual problem. The DI indicator was evident, as the student was able to connect numerical trends to a real-world conclusion, despite limitations in justification depth.

Figure 10 presents ASFZ's conclusion for the same task. ASFZ performed a logical and systematic interpretation, applying an average rate of increase and linear prediction to arrive at a specific projected year. Although the final numerical result differed from the correct answer due to an incorrect starting value, the interpretive process itself was sound and well-reasoned. This was confirmed through interview:

| | | |
|---|---|---|
| *R* | : | *Are you confident your steps are correct?* |
| *ASFZ* | : | *The steps are correct, but maybe the starting number was wrong.* |

This response demonstrates that ASFZ's interpretive process was strong, even though the final result was not fully accurate. The DI indicator emerged clearly, with ASFZ able to connect the

**Table 8**
Summary of Statistical Literacy Indicator Scores Across 37 Field Test Students

| Statictical Literacy Indikator | Average Score (%) | Category |
|---|---|---|
| Problem Understanding | 76,06 | Moderate |
| Data Processing | 69,26 | Moderate |
| Data Interpretation | 68,69 | Moderate |

calculated rate to a contextually meaningful conclusion about Jakarta's projected inundation. Taken together, the contrasting responses of RAM and ASFZ illustrate that the developed task has a positive potential effect in eliciting varying levels of statistical literacy, while also revealing the need for instructional scaffolding to support more rigorous and complete written interpretation.

The analysis conducted during the field test focused on identifying representative patterns of students' responses and illustrating how the developed tasks elicited problem understanding, data processing, and data interpretation.

### Assessment phase

Building upon the field test findings, the assessment phase aimed to examine the overall pattern of students' statistical literacy by analyzing all 37 students' written responses to identify the potential effects of the developed tasks. Consistent with the design research approach, this phase did not aim to measure learning gains, but to analyze how students' reasoning emerged when engaging with the tasks in an authentic classroom setting. Students' written responses were analyzed qualitatively using three indicators of statistical literacy Problem Understanding (PU), Data Processing (DP), and Data Interpretation (DI) treated as analytical categories to capture students' reasoning processes rather than as summative performance measures. The overall results of the analysis are summarized in Table 8.

As shown in Table 8, all three indicators fell within the moderate category, with PU achieving the highest score (76.06%) and DI the lowest (68.69%). Regarding Problem Understanding, most students were able to recognize the problem context, identify relevant information from graphs, tables, and maps, and distinguish between historical and predicted data, although some understood the context only partially by citing numerical values without explaining their meaning in relation to the problem. Regarding Data Processing, students engaged in varied data operations beyond routine calculation, including computing differences between periods, identifying trends, and calculating average rates of change to make predictions, though this engagement was not uniform as some students performed calculations correctly but could not articulate their steps in writing, while others understood the computational logic but made numerical errors. Regarding Data Interpretation, this indicator showed the greatest variation, with some students drawing well-grounded conclusions connected to real-world climate phenomena while others offered overly general or intuition-based interpretations without adequate data support. Overall, these findings indicate that the developed tasks have a positive potential effect in eliciting all three key aspects of statistical literacy, while also revealing the need for instructional scaffolding to support deeper interpretation and more rigorous written justification when working with data involving uncertainty.

## DISCUSSION

The findings of this study indicate that the developed PISA-like task exhibits essential characteristics of a valid and practical assessment instrument within the Uncertainty and Data domain. Task validity was supported by expert judgments confirming the alignment of content, construct, and language with statistical literacy indicators and the intended climate change context, while practicality was reflected in students' ability to comprehend task instructions and engage meaningfully with the data across the one-to-one, small-group, and field test stages. These results are consistent with design research principles that emphasize iterative refinement and empirical grounding of educational products rather than measuring instructional effectiveness (Bakker, 2018; Nusantara et al., 2025; Plomp & Nieveen, 2013; Tanujaya et al., 2023). Moreover, the use of activity-oriented and context-rich tasks aligns with findings by Maphutha et al. (2022), who demonstrated that such tasks support students' engagement and reasoning by promoting active involvement in problem-solving processes rather than procedural execution. In this sense, the developed task functions as a formative assessment instrument capable of eliciting students' reasoning patterns when interacting with complex, data-rich situations (Nugroho et al., 2024; Nusantara et al., 2021; Rahman et al., 2024; Ramli et al., 2022).

The predominance of Problem Understanding and Data Processing over Data Interpretation reflects not merely a developmental limitation of students, but also the interaction between task design characteristics and students' prior assessment experiences. The developed PISA-like tasks intentionally foregrounded rich contextual information and multiple data representations, which effectively supported students in identifying relevant variables, recognizing patterns, and performing basic data processing operations. The climate change context played a specific role in supporting each indicator's emergence. With respect to Problem Understanding, the inherently observable and locally relevant nature of climate change phenomena such as rising temperatures, deforestation, and land subsidence provided students with a meaningful entry point into the data, enabling them to connect the problem situation to real-world conditions they could relate to. This contextual familiarity supported students' ability to identify what the task was asking and to distinguish between historical data and projected values, even when they could not yet express this understanding fully in writing (Prahmana & D'Ambrosio, 2020). With respect to Data Processing, the quantitative richness of climate change data including multi-period percentage changes, trend-based graphs, and comparative tables encouraged students to engage in data operations beyond routine calculation, such as computing rates of change and making data-based predictions, rather than applying a single fixed formula (Islamirta et al., 2022; Nusantara et al., 2021; Nusantara, et al., 2025). Utari et al. (2025b) similarly found that data processing skills develop when students are exposed to contextual data and required to process it step by step. With respect to Data Interpretation, the climate change context provided students with a basis for connecting numerical results to real-world meaning, for instance explaining why a particular region experienced the greatest deforestation or why disaster frequency has trended upward, although many students could articulate these interpretations verbally rather than in written form (Kurnia et al., 2023). This aligns with (Gal, 2002), who argues that interpretation represents a higher-order integration of cognitive and dispositional components requiring not only procedural competence but also critical reflection and contextual reasoning, and with Utari et al. (2025a), who identify interpretation as the most challenging yet most essential aspect of statistical literacy development.

However, transitioning from these processes to coherent data interpretation requires students to articulate relationships between data patterns and contextual meaning, a competence that is rarely emphasized in routine classroom assessments (Ramli et al., 2022; Sujadi et al., 2023). Within the PISA formulate–employ–interpret cycle, the interpret phase demands students to move beyond calculations toward evaluative reasoning, which has consistently been identified as the most challenging stage for lower secondary students (OECD, 2016, 2023). Similar patterns have been documented in studies by Nusantara et al. (2024b) and Prahmana & D'Ambrosio (2020), showing that when students engage with context-based statistical tasks, problem understanding and data processing tend to emerge first, while interpretation develops more gradually and often remains implicit or verbally expressed rather than written. A notable finding in this study is the consistent gap between students' conceptual understanding and their ability to express it in written form across all three indicators, suggesting that the potential effect of the tasks is not limited to final written answers but extends to the reasoning processes revealed through interviews. This condition aligns with findings by Pathuddin et al. (2021) and Nusantara et al. (2024b), who emphasize that students require prior familiarization with context-rich tasks before they can express statistical understanding optimally in writing.

Beyond validity, practicality, and potential effects, this study contributes to the growing literature on the role of meaningful contexts in statistical literacy assessment, particularly through the integration of climate change as a global and data-intensive issue. Climate change data inherently involve variability, long-term trends, and uncertainty, which are central characteristics of the Uncertainty and Data domain (IPCC, 2023; OECD, 2023). The findings indicate that this context supported students' engagement with data and encouraged reasoning based on trends rather than isolated values, extending previous research on context-based PISA-like tasks using socio-scientific issues such as environmental change and public health (Nusantara et al., 2021; Prahmana & D'Ambrosio, 2020). In line with Stacey (2024), the integration of global, data-rich issues into mathematics assessment reflects the increasing need for mathematical competencies that enable learners to interpret complex information in digitally mediated environments. The diversity of strategies and responses produced by students on the same task further confirms the potential effect of the developed instrument, as it allowed varied reasoning pathways rather than constraining students to a single procedure, consistent with findings by Masfingatin et al. (2020). Nevertheless, students' reliance on informal strategies and intuitive interpretations reflects a limitation commonly reported in studies on PISA-like tasks (Gustiningsi et al., 2023; Ramli et al., 2022).

The findings revealed that students were generally able to identify relevant information from graphs and tables and perform basic data-processing procedures, such as comparing temperature changes and recognizing trends in climate datasets. However, many students experienced difficulties when interpreting the results critically and formulating evidence-based conclusions related to the climate change context. In particular, several responses demonstrated that students could extract numerical information correctly but were unable to provide coherent statistical interpretations or justify predictions using the available data. These findings suggest that students' statistical literacy remains stronger at the procedural level than at the interpretative and evaluative levels of reasoning. Therefore, future studies are recommended to integrate climate-based PISA-like tasks into instructional and activity-based learning processes to provide students with more sustained opportunities to engage in data interpretation, argumentation, and evidence-based reasoning. Longitudinal investigations are also needed to examine how continuous exposure to authentic climate-related data can systematically support the development of students' statistical literacy over time, particularly in the domains of data interpretation and reasoning under uncertainty (Büscher, 2022; Kurniadi et al., 2026; Nusantara, et al., 2025).

## CONCLUSIONS

This study employed a design research methodology using Tessmer's formative evaluation to develop a climate-based PISA-like task. The results demonstrate that the task is highly valid (83.9%) and practical (78.8%), effectively engaging students in key statistical literacy components such as problem understanding and data processing. By integrating authentic climate change data, the context served as a meaningful situation for real-world statistical reasoning, although students' interpretations remained largely intuitive. While this study confirms the task's potential as a formative assessment tool, it is limited to qualitative evidence without measuring instructional learning gains. Therefore, future studies should integrate these tasks into structured instructional phases with prior familiarization activities to help students systematically develop and express their statistical reasoning.

## ACKNOWLEDGMENT

This study was partially funded by the DIPA of Universitas Jambi under the Basic Research (Penelitian Dasar) schem. The author gratefully acknowledges the mathematics teacher and the students who participated in this study for their invaluable contributions and active participation.

## AUTHOR'S DECLARATION

| | |
|---|---|
| **Authors' contributions** | DSN contributed to conceptualization, validation, supervision, writing—review and editing, and project administration. SBH contributed to formal analysis, investigation, data curation, and writing—the original draft. FTP contributed to validation, supervision, and methodology. KM contributed to methodology and writing—review and editing. All authors have read and agreed to the published version of the manuscript. |

| | |
|---|---|
| **Funding Statement** | This research was funded by the Basic Research Grant Program of the Institute for Research and Community Service (LPPM), Universitas Jambi, Indonesia. |
| **Availability of data and materials** | All data generated or analyzed during this study are available from the corresponding author upon reasonable request. |
| **Competing interests** | The authors declare that there are no conflicts of interest regarding the publication of this paper. |

## BIBLIOGRAPHY

Aini, I. N., Zulkardi, Putri, R. I. I., & Yaniawati, P. (2022). Developing PISA-like math problems in the content of space and shape through the context of historical buildings. *Journal on Mathematics Education*, *13*(4), 723–738. https://doi.org/10.22342/jme.v13i4.pp723-738

Bakker, A. (2018). *Design Research in Education: A Practical Guide for Early Career Researchers* (1st ed.). Routledge. https://doi.org/10.4324/9780203701010

Britton, M. K., & Anderson, B. A. (2020). Specificity and persistence of statistical learning in distractor suppression. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(3), 324–334. https://doi.org/10.1037/xhp0000718

Büscher, C. (2022). Design principles for developing statistical literacy in middle schools. *Statistics Education Research Journal*, *21*(1), 1–8. https://doi.org/10.52041/serj.v21i1.80

Callaman, R. A., & Itaas, E. C. (2020). Students' mathematics achievement in Mindanao context: A meta-analysis. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *5*(2), 148–159. https://doi.org/10.23917/jramathedu.v5i2.10282

Chick, H. L., & Pierce, R. (2012). Teaching for statistical literacy: Utilising affordances in real-world data. *International Journal of Science and Mathematics Education*, *10*(2), 339–362. https://doi.org/10.1007/s10763-011-9303-2

Deda, Y. N., & Maifa, T. (2021). Development of Student Worksheets Using the Context of Local Wisdom on Integers and Fractions. *Jurnal Pendidikan Matematika*, *15*(1), 71–82. https://doi.org/10.22342/jpm.v.i.12824.71-82

Delport, D. H. (2023). The development of statistical literacy among students: Analyzing messages in media articles with Gal's worry questions. *Teaching Statistics*, *45*(2), 61–68. https://doi.org/10.1111/test.12308

Gal, I. (2002). Statistical Literacy: Conceptual and Instructional Issues. In *Perspectives on Adults Learning Mathematics* (pp. 135–150). Springer. https://doi.org/10.1007/0-306-47221-x_8

Gonda, D., Pavlovičová, G., Ďuriš, V., & Tirpáková, A. (2022). Implementation of pedagogical research into statistical courses to develop students' statistical literacy. *Mathematics*, *10*(11), 1793. https://doi.org/10.3390/math10111793

Gould, R. (2017). Data literacy is statistical literacy. *Statistics Education Research Journal*, *16*(1), 22–25. https://doi.org/10.52041/serj.v16i1.209

Gustiningsi, T., Putri, R. I. I., Zulkardi, & Hapizah. (2023). Developing a PISA-Like mathematical problem: Using traditional food context. *IJECA International Journal of Education & Curriculum Application*, *6*(3), 324–338. https://doi.org/10.31764/ijeca.v6i3.20200

IPCC. (2023). *Climate Change 2021 – The Physical Science Basis*. Cambridge University Press. https://doi.org/10.1017/9781009157896

Islamirta, K. A., Putri, R. I. I., Hapizah, & Mawarni, I. (2022). Development of PISA-Like Math Tasks for Uncertainty and Data Using the Context of COVID-19. *Jurnal Pendidikan Matematika*, *16*(3), 269–288. https://doi.org/10.22342/jpm.16.3.17086.269-288

Johannssen, A., Nataliya, C., Friederike, S., & Stabenow, K. (2021). Statistical literacy—misuse of statistics and its consequences. *Journal of Statistics and Data Science Education*, *29*(1), 54–62. https://doi.org/10.1080/10691898.2020.1860727

Kurnia, A. B., Lowrie, T., & Patahuddin, S. M. (2023). The development of high school students' statistical literacy across grade level. *Mathematics Education Research Journal*, *36*(Suppl 1), S7–S35. https://doi.org/10.1007/s13394-023-00449-x

Kurniadi, E., Zulkardi, Putri, R. I. I., & Darmawijoyo. (2026). Local Instruction Theory: Supporting Derivative Concept Learning Through Climate Change Context. *Mathematics Education Journal*, *20*(1), 27–50. https://doi.org/10.22342/mej.v20i1.pp27-50

Maphutha, K., Maoto, S., & Kibirige, I. (2022). The effect of the activity-based approach on grade 11 learners' performance in solving two-dimensional trigonometric problems. *Eurasia Journal of Mathematics, Science* *and Technology Education*, *18*(10). https://doi.org/10.29333/ejmste/12405

Maphutha, K., Maoto, S., & Mutodi, P. (2023). Exploring grade 11 learners' mathematical connections when solving two-dimensional trigonometric problems in an activity-based learning environment. *Journal on Mathematics Education*, *14*(2), 293–310. https://doi.org/10.22342/jme.v14i2.pp293-310

Masfingatin, T., Murtafiah, W., & Maharani, S. (2020). Exploration of Creative Mathematical Reasoning in Solving Geometric Problems. *Jurnal Pendidikan Matematika*, *14*(2), 155–168. https://doi.org/10.22342/jpm.14.2.7654.155-168

Monteiro, C. E. F., & Carvalho, R. N. (2023). Toward statistical literacy to critically approach big data in mathematics education BT - Research on reasoning with data and statistical thinking: International perspectives. In G. F. Burrill, L. de Oliveria Souza, & E. Reston (Eds.), *Research on reasoning with data and statistical thinking: International perspectives* (pp. 227–242). Springer. https://doi.org/10.1007/978-3-031-29459-4_18

Murdaningsih, S., & Murtiyasa, B. (2016). An analysis on eight grade mathematics textbook of new Indonesian curriculum (K-13) based on PISA's framework. *Journal of Research and Advances in Mathematics Education*, *1*(1). https://doi.org/10.23917/jramathedu.v1i1.1780

Nugroho, A. A., Dwijayanti, I., & Setyawati, R. D. (2024). Development of interactive e-books to increase learning independence in statistics. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *8*(3), 180–189. https://doi.org/10.23917/jramathedu.v8i3.5120

Nusantara, D. S., Zulkardi, & Putri, R. I. I. (2021). Designing PISA-like mathematics problem using a COVID-19 transmission map context. *AIP Conference Proceedings*, *2438*. https://doi.org/10.1063/5.0071596

Nusantara, D. S., Zulkardi, & Putri, R. I. I. (2024a). How to design PISA-like digital mathematics problems: a preliminary study. *AIP Conference Proceedings*, *3046*(1). https://doi.org/10.1063/5.0194756

Nusantara, D. S., Zulkardi, & Putri, R. I. I. (2024b). Students' strategies in solving PISA-like mathematics problems on change and relationship content. *AIP Conference Proceedings*, *3052*(1). https://doi.org/10.1063/5.0201003

Nusantara, D. S., Zulkardi, Z., & Putri, R. I. I. (2025). Investigating Students' Mathematical Argumentation Ability in Solving PISA-like Mathematics Tasks on Change and Relationships. *Jurnal Pendidikan Matematika*, *19*(4), 807–826. https://doi.org/10.22342/mej.v19i4.pp807-826

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education* (Vol. 1). OECD Publishing. https://doi.org/10.1787/9789264266490-en

OECD. (2023). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education (PISA)* (Vol. 1). OECD Publishing. https://doi.org/10.1787/53f23881-en

Pathuddin, H., Kamariah, & Nawawi, M. I. (2021). Buginese Ethnomathematics: Barongko Cake Explorations as Mathematics Learning Resources. *Journal on Mathematics Education*, *12*(2), 295–312. https://doi.org/10.22342/jme.12.2.12695.295-312

Plomp, T., & Nieveen, N. (2013). *Educational Design Research: Part A – An Introduction*. SLO (Netherlands Institute for Curriculum Development). https://www.slo.nl/@4315/educational-design/

Prahmana, R. C. I., & D'Ambrosio, U. (2020). Learning Geometry and Values from Patterns: Ethnomathematics on the Batik Patterns of Yogyakarta, Indonesia. *Journal on Mathematics Education*, *11*(3), 439–456. https://doi.org/10.22342/jme.11.3.12949.439-456

Rahman, F., Juniati, D., & Siswono, T. Y. E. (2024). Covariational reasoning of field dependent prospective mathematics teacher in solving covariance problem. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *8*(3), 140–151. https://doi.org/10.23917/jramathedu.v8i3.4740

Ramli, M., Susanti, B. H., & Yohana, M. P. (2022). Indonesian students' scientific literacy in Islamic Junior High School. *International Journal of STEM Education for Sustainability*, *2*(1), 53–65. https://doi.org/10.53889/ijses.v2i1.33

Rodriguez-Martinez, L. Y., Hernandez-Martinez, P., & Perez-Martinez, M. G. (2023). Development of a protocol to measure mathematics higher-order thinking skills in Mexican primary schools. *Journal on Mathematics Education*, *14*(4), 781–796. https://doi.org/10.22342/jme.v14i4.pp781-796

Schield, M. (2011). Statistical literacy: A new mission for data producers. *Statistical Journal of the IAOS*, *27*(3–4), 173–183. https://doi.org/10.3233/SJI-2011-0732

Sihlangu, S., Maphutha, K., & Mokwana, L. (2025). A commognitive analysis of learners' mathematical thinking on mathematics vocabulary used during classroom discourse. *Journal on Mathematics Education*, *16*(4), 1389–1406. https://doi.org/10.22342/jme.v16i4.pp1389-1406

Stacey, K. (2023). Mathematical competencies in the digital era. *JRAMathEdu (Journal of Research and Advances in Mathematics Education)*, *8*(4), 225–232. https://doi.org/10.23917/jramathedu.v8i4.3096

Sujadi, I., Budiyono, Kurniawati, I., Wulandari, A. N., Andriatna, R., & Puteri, H. A. (2023). The Abilities of Junior High School Students in Solving PISA-Like Mathematical Problems on Uncertainty and Data Contents. *Jurnal Penelitian Dan Pengembangan Pendidikan*, *7*(1), 102–109.

https://doi.org/10.23887/jppp.v7i1.51931

Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2023). Lesson study with sharing and jumping tasks in online mathematics classrooms for rural area students. *Journal on Mathematics Education*, *14*(1), 169–188. https://doi.org/10.22342/jme.v14i1.pp169-188

Tessmer, M. (1993). *Planning and Conducting Formative Evaluations*. Routledge. https://doi.org/10.4324/9780203061978

Ulia, N., Akhsani, L., & Untarti, R. (2023). Statistical literacy process of prospective mathematics teachers: A case study of PISA model problems. *Journal of Higher Education Theory & Practice*, *23*(7), 45–58. https://doi.org/10.33423/jhetp.v23i7.6011

Ünal, D., Çiğşar, B., Alam, D., Atalar, Ş. E., Gümüş, S. İ., & Çakitli, A. (2023). Reading the world with statistical literacy: Results of an emprical study. *Bilge International Journal of Science and Technology Research*, *7*(2), 198–203. https://doi.org/10.30516/bilgesci.1251429

Utari, R. S., Putri, R. I. I., & Zulkardi. (2024). Designing a Hypothetical Learning Trajectory using the Local Wisdom of South Sumatera as a Context through Hybrid Learning. *Jurnal Pendidikan Matematika*, *18*(1), 79–96. https://doi.org/10.22342/jpm.v18i1.pp79-96

Utari, R. S., Putri, R. I. I., Zulkardi, & Hapizah. (2025a). Advancing statistical literacy for prospective teachers: A systematic review of contemporary approaches and best practices. *Jurnal Elemen*, *11*(2), 390–407. https://doi.org/10.29408/jel.v11i2.28044

Utari, R. S., Putri, R. I. I., Zulkardi, & Hapizah. (2025b). Supporting statistical literacy skills for prospective teachers: A learning trajectory used South Sumatra local wisdom context through hybrid learning. *Infinity Journal*, *14*(3), 711–732. https://doi.org/10.22460/infinity.v14i3.p711-732

van Dijke-Droogers, M., Drijvers, P., & Bakker, A. (2022). Introducing statistical inference: Design of a theoretically and empirically based learning trajectory. *International Journal of Science and Mathematics Education*, *20*(8), 1743–1766. https://doi.org/10.1007/s10763-021-10208-8

Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, *96*(1), 33–47. https://doi.org/10.1007/s10649-017-9764-5

Zulkardi. (2002). *Developing a learning environment on realistic mathematics education for Indonesian student teachers* [University of Twente]. https://doi.org/10.3990/1.9789036518451