Development of Statistical Literacy Questions for Prospective Mathematics Teachers

Rahma Siska Utari\textsuperscript{1}, Zulkardi\textsuperscript{2*}, Ratu Ilma Indra Putri\textsuperscript{3}, Ely Susanti\textsuperscript{4}

\textsuperscript{1}Pendidikan Matematika, Universitas Sjakhyakirti, Indonesia
\textsuperscript{2,3,4}Pendidikan Matematika, Universitas Sriwijaya, Indonesia
\textsuperscript{*}zulkardi@unsri.ac.id

Abstract
Statistical literacy was one of the important abilities that must be possessed by prospective mathematics teachers in the era of society 5.0. This research was qualitative descriptive to develop statistical literacy questions for prospective mathematics teachers using indicators of the five bases of statistical literacy and PMRI principles. The questions were developed using a design research type of development study to see the validity and practicality of the questions. The research stages included the preliminary stage, and the prototyping stage (self-evaluation, expert review, one-to-one, and small group). As many as 13 prospective mathematics teachers participated in this study, in the one-to-one stage and the small group stage. Data were collected through walkthroughs, observations, interviews, and video recordings. Data were analyzed qualitatively to describe the results of each stage. The results showed that the statistical literacy questions developed were valid and practical. Validity was shown from the expert review stage which validates the content, construct and language aspects used in the questions. Practicality can be seen from the one-to-one and small group stages where students can work on the questions given. Further research was needed using more subjects to see the potential effects that emerge from this study.

Keywords: Design Research, Development Study, PMRI, Statistical Literacy

Received: December 05, 2022/ Accepted: January 23, 2023/ Published Online: January 26, 2023
INTRODUCTION

Literacy and numeracy were basic competencies that important competencies in the 21st century that students must possess to be able to compete globally and they were characteristics of the Kurikulum Merdeka (Kemdikbud, 2022b; Pusmendik, 2022; Sinaga, 2022). One of the elements of mathematics content in the curriculum was data and opportunities, they were statistics. Statistics studied data, data types, data representation, data analysis related to data central tendency, dispersion of data, opportunities, and uncertainties (Allen et al., 2020; Kemdikbud, 2022a). Data were part of statistics that were often used for human activity. The importance of using statistics in everyday life and all areas of life increases attention to statistical literacy (Ben-Zvi, 2020; Gal, 2019).

Statistical literacy was important in the digital era where students were continuously presented with data, and statistical information from various sources in society (Aziz & Rosli, 2021; Sharma, 2017). Statistical literacy was the ability to interpret, criticize, evaluate, and communicate statistical information and messages through various forms of media (Aziz & Rosli, 2021; Garfield & Burrill, 1997; Rumsey, 2002). Statistical literacy was the ability to critically interpret and evaluate statistical information in various contexts and to communicate understanding in ways that can impact decision-making (Gal, 2004, 2019; Moore, 1997; Sharma, 2017; Watson, 2011). Prospective mathematics teachers as teachers in the future are required to have good statistical literacy skills to face the digital era along with its challenges (Sinaga, 2022).

Studies showed that the statistical literacy abilities of prospective mathematics teachers in Indonesia were not optimal (Andriatna et al., 2021; Idris, 2019; Khaerunnisa & Pamungkas, 2017; Tiro, 2018). Khaerunnisa & Pamungkas (2017) stated that as many as 52% of students majoring in mathematics education had low statistical literacy skills, students were unable to use relevant information from statistical concepts, and were unable to interpret and conclude statistical problems. Andriatna et al. (2021) stated that as much as 46.67% of the statistical literacy skills of prospective mathematics teachers were not optimal in terms of basic statistical conceptual reasoning indicators, resulting in erroneous justification and interpretation of data. When students made mistakes in interpreting data and information resulting in errors in making conclusions and communicating them in various media (Gal, 2019; Sharma, 2017).

In other words, to optimize and support the statistical literacy skills of prospective mathematics teachers as mathematics teachers in the future, it was necessary to develop statistical literacy instruments, such as statistical literacy questions (Fernández et al., 2020;
The urgency of this research was to support the statistical literacy skills of prospective mathematics teachers by developing questions as their reference for practicing and learning, if they have good statistical literacy knowledge it can help minimize errors in interpreting data and drawing conclusions in real-life problems (Gal, 2019; Takaria & Talakua, 2018; Watson, 2011; Weiland, 2017).

Gal (2004) stated that developing statistical literacy requires joint activation of five interrelated knowledge bases namely: literacy, statistics, mathematics, context, and also critical thinking. When developing contextual statistical literacy and modeling becomes the basis of learning activities, statistics were numbers in context, and the context was a source of meaning and basis for applying statistical procedures/modeling and for interpreting the results obtained (Gal, 2019). The use of context, mathematics, and models in learning corresponded to the principles and characteristics of Realistic Mathematics Education (RME), here in after known in Indonesia as Pendidikan Matematika Realistik Indonesia (PMRI). The context in learning mathematics supports students to acquire a lot of mathematical knowledge, the context bridges informal knowledge to formal mathematics (Wijers & de Haan, 2020; Zulkardi, 2002).

Previous studies, Khaerunnisa & Pamungkas (2017) and Andriatna et al. (2021) did not focus on the use of context in developing students' statistical literacy skills. Idris (2019) described the perspective of lecturers and students that the use of cultural contexts and Islamic values can be used for statistical literacy. The novelty of this study was to develop statistical literacy questions that focused on five knowledge bases including literacy, statistics, mathematics, context, and critical thinking, and integrated with PMRI principles, namely the use of contexts and models which had not been done in previous research. This research was a preliminary study in developing initial prototypes of statistical literacy questions for prospective mathematics teachers using a socio-cultural context. Follow-up research was conducted to see the potential effects arising from the development of statistical literacy questions. The purpose of this study was to develop statistical literacy questions for prospective mathematics teachers using the five bases of statistical literacy and PMRI principles that are valid and practical.

**METHOD**

This research was qualitative descriptive research to describe the process of developing statistical literacy questions with a socio-cultural context using a design research type of development study. The stages used are the prototyping (formative evaluation) stage which...
consists of self-evaluation, expert review, one-to-one, small group, and field tests. As shown in Figure 1. The following flowchart was for the development of statistical literacy questions.

![Flowchart of the development of statistical literacy questions](image)

**Figure 1.** Flowchart of the development of statistical literacy questions (Tessmer, 1993; Zulkardi, 2002)

Based on Figure 1. Flowcharts of the development of statistical literacy questions, in this study to see the validity and practicality of statistical literacy questions, were carried out up to the small group stage. As many as 13 students of the mathematics education study program from Sjakhyakirti University participated in this study. Data was collected using tests, observations, interviews, walkthroughs, and video recordings. The data were analyzed qualitatively to see the validity and practicality of the questions developed.

Walkthroughs were used to collect data in the form of suggestions, opinions, and comments from expert reviews to determine the validity of the questions based on content, construct, and language. The questions were valid if the content in the question meets the five statistical literacy bases and the six PMRI principles. The following was Table 1. The five statistical literacy bases and the Six PMRI Principles

**Table 1.** Five statistical literacy bases and six PMRI principles

<table>
<thead>
<tr>
<th>Statistical literacy base</th>
<th>PMRI principle</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Literacy skills (literacy skills) in the form of skills to examine statistical messages conveyed either orally or in written text.</td>
<td>• The activity principle was that students are treated as active participants in learning, mathematics is learned through mathematics and education.</td>
</tr>
<tr>
<td>• Statistical knowledge can be in the form of basic statistical ideas with graphic or table representations and their interpretations, knowing the basic notation of opportunities, and</td>
<td>• The reality principle where the purpose of learning mathematics is to apply mathematics in solving real-life problems.</td>
</tr>
</tbody>
</table>
The questions were developed using indicators in Table 1. Five statistical literacy bases and six PMRI principles. From the constructed aspect, the questions were valid if it corresponds to the competencies of the basic statistics course and the existing curriculum. In the language aspect the questions were valid if they question is by the *Pedoman Umum Ejaan Bahasa Indonesia* (PUEBI), the language used in the questions was not ambiguous and contained multiple interpretations. The problems were practical if the question can be done by students at the one-to-one and small group stages.

**RESULT**

The preliminary stage was carried out to get an overview of statistical literacy and PMRI. At this stage, a literature review was carried out regarding the competence of basic statistics courses in higher education, especially in mathematics education study programs, the basis of statistical literacy, PMRI, and their interrelationships. The combination of basic statistics subject competencies, statistical literacy basis, and PMRI was used as an indicator for developing statistical literacy questions. Statistical literacy questions are designed at this stage.
Questions that have been designed at the preliminary stage are continued at the formative evaluation stage or the prototyping stage.

The first step in the formative evaluation was self-evaluation, at this stage, the researcher reviews the process and produces statistical literacy questions and the scoring rubric. The results of the self-evaluation stage were referred to as prototype I. Figure 2 below was a matter of the results of prototype 1.

![Figure 2](image_url)

**Figure 2.** Designing problems and statistical literacy questions (prototype 1)

**Figure 2** Designing problems and statistical literacy questions were the results of the prototype 1 process, the questions were designed using indicators of the five statistical literacy bases namely, (a) literacy skills, skills to read messages or study statistical messages from readings in the form of news headlines, (b) knowledge statistics, were statistical ideas and concepts used such as measures of data concentration, in the form of averages, data representation in various forms, estimates in the form of opportunity theory based on social situations, current society, (c) mathematical knowledge can be in the form of using addition number operations, multiplication, estimation and also included other mathematical concepts that are used to support other statistical knowledge, (d) context knowledge, namely socio-
cultural news headlines in the community in the form of LRT public transportation users which were the source of the meaning and basis of interpretation of the results of data analysis, as well as (e) critical questions asked expect emerging from reading results to critically evaluate statistical information. In addition to using the five statistical literacy bases, the questions are also designed based on PMRI principles.

The results of prototype 1 were reviewed by experts to see the validity of the questions from the aspects of content, construction, and language, this stage was called the expert review stage. For the expert review stage, two lecturers in the mathematics education study program participated in this study, and an interview process was conducted with the two lecturers. As the expert review progressed, the questions were also tested in the one-to-one stage involving three students from the education study program, mathematics. The one-to-one stage was carried out to see how students can understand or understand the meaning of the questions and answer questions. At this stage, interviews were also conducted with students to get comments and suggestions on how students understand and solve problems. Comments and decision results from the expert review and one-to-one stages were described in Table 2. Decision results from prototype 1.

Table 2. Comments and decision results from prototype 1

<table>
<thead>
<tr>
<th>No</th>
<th>Validator</th>
<th>Comment</th>
<th>Decision</th>
</tr>
</thead>
</table>
| 1  | RR        | ● The news text was too long, was it possible to re-narrate the headline?  
● The problems presented already contain five statistical literacy bases, the expected statistical competencies were by the curriculum, the language used was by PUEBI. | ● Design the display of news headlines and questions to be laid out again so that it was more attractive and not too long  
● Questions from news headlines are still used. |
| 2  | LA        | ● In terms of layout (view) it’s better to use applications that are widely used today (Canva, etc.) so that the display was more attractive and informative  
● From the aspect of language, the meaning of the questions was easy to understand and does not cause double meanings, five statistical literacy bases have also been raised, and the context also exists. | |
Table 2. The comments and decision results from prototype 1 were the reference table used to revise prototype 1 so that the results from prototype 1 which have been revised become prototype 2. Prototype 2 was tested at the small group stage which aims to see the practicality of the questions being developed. In the small group stage, as many as 10 students of the mathematics education study program participated. 10 students were divided into two groups, and students discussed with each other in groups to solve the problems given. At this stage, observations and interviews were carried out while students were solving the given problems. Interviews were conducted to see how students' thinking processes understood the problems presented and how students' understanding of solving statistical literacy problems was given. Interviews between researchers and students in the small group stage were described in Transcript 1. The conversation at the small group stage was the estimation of LRT users in 2022 below.

Transcript 1. Conversation at the small group stage, estimation of LRT passengers in 2022

Researcher : "Did you have trouble estimating the number of LRT passengers in 2022?"

Student 4 : "So far it's safe ma'am, this has got the estimation results. We estimate the total number of LRT passengers in 2022, using data in news 2."

Researcher : "How's the process going? can be explained? Titled in news 2 only information is given on LRT users until August 2022."

Student 5 : "For 8 months in 2022, you will get 1.79 million passengers, then divide 1.79 million by 2 to get 895,000 passengers for 4 months. So the estimate until December 2022 is 895,000 multiplied by 3. So the estimated LRT users until December 2022 are 2,685,000 passengers."

Researcher : "How about the other guys? If I divide 1.79 million by 8, was that OK?"

Student 6 : "You can ma'am, if divided by 8 means later the result will be multiplied by 12 ma'am, we will calculate the number of passengers for a month"
Transcript 1. The conversation at the small group stage of estimation of LRT passengers in 2022 above explained that from the five knowledge bases of statistical literacy and PMRI principles used, students had used literacy skills, namely, students can already study statistical messages in writing on news headlines. So that from this message they can use available information to estimate the number of LRT passengers in 2022 by using statistical ideas (statistical knowledge) in the form of an average concept, for mathematical knowledge that emerges there was a concept of comparative reasoning that supports the concept of statistical average so that students can estimate, the socio-cultural context using the Palembang city public transportation context was a source for interpreting the estimation results of LRT passengers up to December 2022. Critical questions were needed to be able to evaluate the results of the estimation of the number of LRT passengers in 2022. Group answers in writing can be seen in Figure 3. Student answers in estimating the number of LRT passengers.

![Figure 3](image.png)

**Figure 3.** Student answers in estimating the number of LRT Passengers

Figure 3. Students' answers in estimating the number of LRT passengers explained that students have used the basic concept (statistical idea) in the form of an average of four months of LRT passengers to solve problems and also the mathematical comparison concept to find an estimate of the overall number of passengers in 2022. The next problem namely in representing LRT user data from 2018 to 2022 can be seen in Figure 4. Student answers in representing LRT user data.
Figure 4. Student answers in representing LRT passengers’ data

Figure 4 student answers in representing LRT passengers data showed that students represent data using tables as well as line charts. The line chart above also shows that LRT passengers data from 2018 fluctuated or experienced increases and decreases, in 2020 due to the Covid-19 pandemic the number of passengers decreased dramatically, 2021 was the new normal era so LRT passengers will gradually increase and in 2022 the estimated number of passengers reached 2,685 million passengers, not too different for the passengers’ data in 2019.

The next question was student interpretation regarding the prediction of the number of LRT passengers in 2023 by considering all social, cultural, and economic aspects in a society that has moved dynamically in the last 4 years. Students presented the results of their discussions in small groups in front of the class. In the explanation process, the researcher conducted discussions and questions and answers (interviews) to explore students' thought processes in interpreting information from data and making conclusions. The following was Figure 5. Students presented the results of the discussion in small groups.
Figure 5. Students presented the results of discussions in small groups

Figure 5. Students presented the results of discussions in small groups to answer the three questions posed, after the presentation, there was a question and answer session to see students' thought processes in interpreting existing data to predict and make conclusions from the message. The discussion process which was also an interview process can be seen in Transcript 2. Student conversations in interpreting data.

**Transcript 2.** Student conversation in interpreting data

**Researcher**: "How are your answers related to the number of LRT users in 2023?"

**Student 7**: "Looking at the passenger chart from 2018 to 2022, we see that the predictions for 2023 passengers will not be much different from those for 2022."

**Researcher**: "Can another friend explain further?"

**Student 8**: "So it's like this ma'am, the graph can go up or down but it won't be too far from the 2.6 million figure. Because seeing the current situation where people are doing their activities as usual, after the new normal due to Covid, and the data shows that the number of passengers until August 2022 is gradually almost the same as normal passengers in 2019 touching 2.6 million."

**Student 9**: "Slightly added ma'am, fuel prices that will increase in 2022 may also result in the number of passengers being stable or not decreasing, because ticket prices are affordable and the facilities are comfortable from buying expensive pertalite"

**Researcher**: "Do you think there is a possibility that the number of passengers will decrease as drastically as in 2022?"

**Student 7**: "It's possible ma'am, it really could happen, the number of passengers is small, like the Covid-19 case which is very sudden and makes a change in society, but if you look at the current situation, and hopefully there won't be an endemic or pandemic in the coming year, we think that the graph will not be far from 2019 and 2022."

**Researcher**: "Ok, pretty clear why."
Transcript 2. Student conversations in interpreting data can be seen in how students criticize information and phenomena that occur in society and what they experience in everyday life so that they can make decisions /conclusions to predict the results of LRT passengers in 2023 by utilizing available data sources. exist in texts, information, and personal experiences, phenomena that exist in society, and make the answers supported by existing data. The small group stage was a stage that is used to see the practicality of the developed statistical literacy questions.

DISCUSSION

The development of statistical literacy questions using a socio-cultural context in the form of the context of news headlines on LRT passengers helped students build information from the given context. The context of LRT passengers was a context that contains problems that exist in everyday life, the use of LRT passenger contexts to support statistical literacy skills corresponds to context knowledge and reality principles and level principles in PMRI, According to Gal (2004) context knowledge was the basis for interpreting results analysis and interpretation of data, according to this Heuvel-panhuizen et al. (2014) stated that the use of context corresponds to the level principle in PMRI and the reality principle. The level principle stated that informal contexts help students build insight and determine concepts and strategies that can be used to solve problems and the reality principle states that the goal of learning statistical literacy (part of mathematics) was to apply mathematics in solving real-life problems.

In addition to using the socio-cultural context, good literacy skills from reading news texts were needed so that there was no misinterpretation of reading results which results in the problem-solving process, both the selection of solving strategies and knowledge of mathematics, statistical knowledge and the necessary critical questions. Transcript 1 and Transcript 2, showed that good literacy skills, supported by mathematical knowledge, statistical knowledge, and critical questions make students' thinking processes develop by utilizing mathematical knowledge, phenomena that exist in society, so they can solve existing problems. Sharma (2017) stated that with statistical literacy students can predict and choose strategies for solving existing problems by utilizing statistical abilities, statistical competence, and information that exist in society. In PMRI, a situation like this was consistent with the intertwinement principle, which offered rich problems where various tools and mathematical knowledge can be used. The guiding principle was that educational programs must contain scenarios that have the potential to work as levers in achieving changes in student comprehension.
Joint problem-solving activities in groups were also consistent with PMRI principles, namely the activity principle where students were treated as active participants in learning and the interactive principle signifies that learning mathematics is not an individual activity but a social activity, by working in groups socio-mathematics was formed in learning which helps students enrich their thinking processes. explaining thoughts with explanations, proofs, and justifications (Utari, 2017; Yackel, 2001; Yackel & Cobb, 1996).

In this study students were able to use relevant information by using mathematical and statistical concepts, to solve problems so that in interpreting the data to predict the possibility they had used existing data, it can be said that the questions designed and developed can support the statistical literacy abilities of prospective mathematics teachers. This research was different from the results of previous studies which stated that students were unable to use relevant information and statistical concepts and made mistakes in interpreting data. Furthermore, Andriatna et al. (2021) and Khaerunnisa & Pamungkas (2017) state that errors in interpreting data and information result in errors in making conclusions and communicating them in various media. The design of statistical literacy questions that use indicators of the five bases of statistical literacy and PMRI principles helps to support the statistical literacy abilities of prospective mathematics teachers.

**CONCLUSION**

From the results of the research and discussion, it can be concluded that the development of statistical literacy questions for prospective mathematics teachers using the five bases of statistical literacy and PMRI principles was declared valid and practical. The validity of the content, constructs, and language was seen from the expert review process which stated that the questions developed had fulfilled the characteristics of the statistical literacy base and PMRI principles, based on the applicable basic statistics course curriculum and the language used already used PEUBI. The practicality of the statistical literacy questions developed was seen from the one-to-one and small group stages, seen from comments, and written and oral answers from students which showed that the statistical literacy questions developed were practical to use and support students in solving the questions given to developing statistical literacy skills.

This study needs to be carried out to develop questions with statistical literacy ability indicators which are important to support future mathematics teachers, and contribute as a reference for future mathematics teachers, researchers, and others, where there was currently not much research on statistical literacy. Further research is needed using more subjects to see the potential effects that emerge from this study.
REFERENCES


