Data literacy of high school students on physics learning

Suryadi\textsuperscript{1,2}, I K Mahardika\textsuperscript{1}, Supeno\textsuperscript{1} and Sudarti\textsuperscript{1}

\textsuperscript{1}Science Education Study Program, University of Jember, Indonesia
\textsuperscript{2}SMA Negeri Jenggawah Kabupaten Jember, Jawa Timur, Indonesia

E-mail: supeno.fkip@unej.ac.id, zuryathe@gmail.com

Abstract. Data literacy is the ability to collect, understand, manipulate, and use data. Based on the viewpoint of cognitive skills, data literacy is related to the ability to collect, filter, select, analyse, interpret, critique, visualize, and communicate data. This ability is essential for students to have, especially in solving learning physics and everyday life problems. Data literacy is indispensable in the life of the modern world like today. Students who have good data literacy will understand and assess data critically, whether presented in tables, graphs, or charts. The purpose of this study is to determine the data literacy of high school students in physics learning. Data were collected through the measurement of data literacy using written tests. Measured data literacy included students' ability to collect, analyse, interpret, implement and evaluate data as well as assess data quality. The results show that high school students' data literacy in physics learning is still in the unsatisfactory category. Various efforts must be made so that data literacy for high school students can be develop, i.e., by implementing learning strategies and using appropriate learning resources. Another action that can be done is to apply a teaching model that follows the characteristics of 21st-century learning.

1. Introduction

The global world is developing very rapidly, so it demands various new abilities, including data literacy. The term data literacy is widely used to describe the ability to use data as part of thinking and reasoning activities to solve real problems [1] and the ability to make decisions [2,3,4]. Data literacy includes the ability to formulate and answer questions using data as part of evidence-based thinking; use appropriate data and representations to support ideas; interpreting information based on existing data; developing and evaluating data-based inferences and explanations, and use data to solve real problems and communicate them to others [5]. Data literacy includes knowledge and skills in assessing, collecting, and analysing data to test a research hypothesis [6] so that it becomes like a part of the science process and higher-order thinking skills [7,8]. Data literacy-related the ability to construct scientific explanations based on data [1], which can be obtained through the inquiry process [9].

Data literacy is a life skill that must be developed because current problems are often related to data. Individuals are often faced with the problem of evaluating and assessing a phenomenon and making decisions using data. Data literacy is the skill and knowledge needed in school education [10]. Data literacy is necessary for students to investigate authentic problems [11]. Data literacy is needed for students as evidence to support scientific reasoning and explanation [12]. Students who are familiar with data processing activities have learned to solve problems [13]. Several research results show that
students who are used to collecting, analyzing, and interpreting data can construct evidence-based scientific explanation and reasoning [14,15,16,17,18].

Data literacy also provides benefits for teachers, both in school administration management and in the learning process. Data literacy helps teachers carry out a consistent and systematic assessment process and supports the decision-making process [4] to provide an overview of student learning outcomes. The method of improving the assessment carried out by the teacher can be used to improve the quality of learning. Data literacy can be utilized by teachers in learning practices [19], allowing teachers to deepen learning objectives and increase students’ feedback [20,21] so that it impacts the learning process and improves students’ learning outcomes [22,23].

Although data literacy provides many benefits for students, the existing conditions show that students often have difficulty obtaining, managing, and interpreting data to support their learning process. Students have problems writing data collection procedures and analyzing data that has been obtained from investigative activities [7]. Students must work hard to get relevant data and interpret it [24]. Some students have minimal knowledge and skills about data [25,26,27,28]. The difficulties that are often experienced related to data generally are in the process of finding, understanding, manipulating, and using data [29].

The results show that students and teachers need data literacy, but the facts show that the students’ ability to obtain, analyse, and interpret data is still unsatisfactory. Does this condition also occur in physics learning in high school? Thus, it is necessary to conduct research to obtain an overview of high school students’ data literacy in physics learning. The research results are expected to be used as a reference for physics teachers in formulating learning objectives, designing learning tools, determining applied learning strategies, designing learning resources, and conducting assessments oriented towards developing data literacy for high school students in physics learning.

2. Method
This research is included in the form of a descriptive study designed to document the condition of high school students' data literacy after learning physics. Data literacy is described without considering the treatment given during the physics learning process. Data were collected on 60 students from 5 state high schools in Jember, East Java Province, Indonesia, namely SMAN 1 Jember, SMAN Arjasa, SMAN Sukowono, SMAN Ambulu, and SMAN Jenggawah. In each school, 12 students were selected based on the physics learning outcomes that have been obtained, consisting of 3 students with good abilities, three students with moderate skills, and three students with low skills. Data literacy was measured on the 60 students using a written essay test of 6 items, according to the aspects contained in data literacy. The measured data literacy aspects included collecting data, accessing data quality, analyzing data, interpreting data, implementing data, and evaluating data. The description for each component of data literacy refers to research conducted by [6].

The results of the data literacy test were scored based on the scoring guidelines made by the researcher. Student scores are categorized into three categories, score 1 for not answering or answering wrong, score 2 for answering correctly but the incomplete description, 3 for answering correctly with complete description.

| Table 1. Aspects of data literacy and its description. |
|------------------------------------------------------|
| Aspects     | Description                                                        |
| Collect data | Recording each of the five data from the activities carried out related to physics material |
| Assess data quality      | Retrieving data that can be used to determine quantities in physics learning |
| Analyse data   | Determining the quantity specified in physics learning             |
Interpret data Delivering the relationship between the data obtained, resulting in the calculation of the specified quantity
Implement data Using the results for the use of everyday life
Evaluate data Things that make data literacy easy/difficult

The data literacy test scores obtained by each student at each school were tabulated for descriptive statistical analysis. Descriptive statistical analysis was performed to determine the highest score, lowest score, average score, standard deviation, and the percentage of students for each level of the score achieved. The subsequent data analysis is to determine the number of students for each achievement of the data literacy score level. This analysis is carried out for all aspects of data literacy assessment, including collecting data, analysing data, interpreting data, implementing data, evaluating data and data quality access. The results of the analysis obtained were then tabulated and interpreted verbally.

3. Result
The data literacy assessment results were carried out by referring to the scoring guidelines developed by the researcher. The achievement of each student's score, which was initially on a scale of 1-18, was converted into data literacy scores on 0-100. The analysis of students' data literacy scores in physics learning is shown in Table 1. Student data literacy from five SMA in Jember Regency is tabulated based on the achievement of the lowest score, the highest score, the average, and the standard deviation.

Table 2. Summary of data literacy descriptive statistics.

| Senior High Schools | Highest Achievement (scale 0-100) | Lowest Achievement (scale 0-100) | Average | Deviation Standard |
|---------------------|----------------------------------|----------------------------------|---------|-------------------|
| SMAN 1 Jember        | 100                              | 44                               | 70.55   | 20.63             |
| SMAN Arjasa          | 89                               | 33                               | 60.00   | 19.56             |
| SMAN Sukowono        | 78                               | 33                               | 47.22   | 13.42             |
| SMAN Ambulu          | 100                              | 33                               | 63.89   | 22.57             |
| SMAN Jenggawah       | 89                               | 33                               | 58.89   | 18.56             |

Even though the highest score reaches the maximum value, the average data literacy score for high school students in learning physics is unsatisfactory. Likewise, the lowest data literacy score achieved by students in all schools is in a low category. With the achievement of the lowest score, it can be said that some students are still unable to give correct answers. Of the five senior high schools, only students of SMAN 1 Jember achieved an average above 70, students from four other high schools received an average of below 70. Based on the data in Table 1, it can be seen that SMAN 1 Jember has the average data literacy. The score is the best compared to other schools. The data literacy skills of students in these schools have been developed even though they are not optimal. In the other four schools, the average score is still in the unsatisfactory category. This result shows that the data literacy of students in these schools is not optimal.
Table 3. Distribution of the number of students for each level of data literacy.

| The aspect of data literacy | Score level | Number of students | Percentage |
|----------------------------|-------------|--------------------|------------|
| collect data               | High        | 16                 | 32%        |
|                            | Middle      | 26                 | 52%        |
|                            | Low         | 8                  | 16%        |
| assess data quality        | High        | 3                  | 6%         |
|                            | Middle      | 26                 | 52%        |
|                            | Low         | 21                 | 42%        |
| analyze data               | High        | 14                 | 28%        |
|                            | Middle      | 26                 | 52%        |
|                            | Low         | 10                 | 20%        |
| interpret data             | High        | 7                  | 14%        |
|                            | Middle      | 20                 | 40%        |
|                            | Low         | 23                 | 46%        |
| implement data             | High        | 6                  | 12%        |
|                            | Middle      | 19                 | 38%        |
|                            | Low         | 25                 | 50%        |
| evaluate data              | High        | 7                  | 14%        |
|                            | Middle      | 18                 | 36%        |
|                            | Low         | 23                 | 46%        |

The results of data literacy analysis for each aspect are shown in Table 2. Data literacy in each aspect is grouped into three categories, high, medium, and low. The distribution of the number of students based on the level of achievement scores for each aspect of data literacy is shown in Table 2.

Based on the data in Table 2, most students' data literacy in the aspect of collecting data is included in the high and medium levels. This result shows that the students have been able to carry out measurement activities on physical quantities and collect data. However, students have not been able to choose the appropriate data with physics problems. This result indicates that most students' scores in assessing the quality of the data are at the middle and low levels. Most of the students' data literacy for data analysis aspects are at a moderate level. This result shows that students are not used to using data in performing mathematical calculations using physics formulations. Most of the data literacy for the implementation and evaluation of data is at medium and low levels. This result shows that students have not been able to use the information that has been obtained in the context of different problems.

In general, the results show that the data literacy of high school students in physics learning is in the unsatisfactory category. Students are less capable of accessing data quality, analysing, interpreting, implementing and evaluating data. The results of this study are identical to previous studies that students often failed to organize or write about the data in any coherent way and treated all data as equally important [30]. The results of this research are not following the characteristics of the material and physics learning. In terms of content, physics material is always related to various representations; symbols, verbal, pictures, graphics, tables and mathematics. Some of these representations are associated with physical quantity measurement data. Although various physics representations are always taught to students, it is not certain that data literacy has become part of students' activities during the physics learning process. By considering the various use of data literacy and the results of this study, it is necessary to integrate multiple learning activities that can develop students' ability to collect data, access data quality, analyse, interpret, implement, and evaluate data. This is in line with what is suggested by [1] that integrating data literacy in learning as part of the curriculum in schools is an essential aspect of ensuring that graduates will have the necessary skills in data. Students can be
trained to carry out a series investigation processes, from planning activities, collecting data, analyzing data, interpreting data, and formulating conclusions. With these activities, students are used to dealing with data.

4. Conclusion
Based on the results of this study, it can be concluded that high school students' data literacy in physics learning is in the unsatisfactory category. Although students can already collect data, most students are still unable to access data quality, analyze data, interpret, implement and evaluate data. Physics subject matter, which is always related to data, has not made students aware of the abundance of data that can be used around it. Various forms of data as representations of real physics phenomena must be applied in solving problems in everyday life. Students can analyze multiple phenomena in the physics learning process but still have difficulty solving problems related to daily life.

Physics learning must be able to teach students about data literacy. Students must be accustomed to representing various physical phenomena in the form of data. Therefore, it is necessary to have the right learning environment and learning resources to learn about collecting, processing, and analyzing data. One alternative that physics teachers can do to teach data literacy is to implement appropriate approaches, models, and methods to implement learning strategies. Teaching materials and worksheets must integrate data literacy so that students can carry out learning activities, both cognitively and physically. Data literacy can become a habit that students can do by observing real physical objects or phenomena around them, making it easier to equip physics learning and solve everyday life problems.

Acknowledgement
We would like to thanks for the support from Faculty of Teacher Training and Education University of Jember.

References
[1] Wolff, A, Gooch, D, Cavero Montaner, J J, Rashid, U, & Kortuem, G 2016 Creating an understanding of data literacy for a data-driven society The Journal of Community Informatics 12(3) 9-26
[2] Mandinach, E B and Gummer, E. S 2013 A systemic view of implementing data literacy in educator preparation. Educational Researcher 42(1) 30–37
[3] Schildkamp, K, Karbautzki, L, & Vanhoof, J, 2014 Exploring data use practices around Europe: Identifying enablers and barriers. Studies in Educational Evaluation 42 15–24
[4] Reeves, T D and Honigh S L 2015 A classroom data literacy intervention for pre-service teachers. Teaching and Teacher Education 50 90-101
[5] Vahey, P, Yarnall, L, Patton, C, Zalles, D, & Swan, K 2006 Mathematizing middle school: Results from a cross-disciplinary study of data literacy Annual Meeting of the American Educational Research Association, San Francisco, CA
[6] Ebbeler, J, Poortman, C L, Schildkamp, K, & Pieters, J M 2017 The effect of a data use intervention on educators' satisfaction and data literacy Educational Assessment, Evaluation and Accountability 29(1) 83–105
[7] Supeno, Astutik, S, Bektiarso, S, Lesmono, A D, & Nuraini, L 2019 What can students show about higher order thinking skills in physics learning? IOP Conference Series: Earth and Environmental Science 243(1) 12127
[8] Supeno, Astutik, S, dan Lesmono, A D 2020 Pre-service science teachers' understanding of scientific method for studying local environmental issues. IOP Conference Series: Earth and Environmental Science 485(1), 012033. doi:10.1088/1755-1315/485/1/012033
[9] Williams, S, Deahl, E, Rubel, L, & Lim, V 2014 City digits: Local lotto: Developing youth data literacy by investigating the lottery Journal of Digital Media Literacy
[10] Gummer, E S and Mandinach, E B 2015 Building a conceptual framework for data literacy Teachers College Record 117(4)
[11] Cobb, G, and Moore, D 1997 Mathematics, statistics, and teaching The American Mathematical Monthly 104 801–823
[12] Rubin, A 2005 Math that matters Hands On: A Journal for Mathematics and Science Educators 28(1) 3–7
[13] Erwin, R 2015 Data literacy: real-world learning through problem-solving with data sets American Secondary Education 43(2), 18–26
[14] Anjani, F, Supeno, Subiki 2020 Kemampuan penalaran ilmiah siswa SMA dalam pembelajaran fisika menggunakan model inkui terbimbing disertai diagram berpikir multidimensi. Lantanida Journal 8(1) 13-28
[15] Imaniar, B. O, Supeno, dan Lesmono, A D 2020 Argumentation of senior high school students on physics instruction based inquiry Compton 7(1) 35-47
[16] Prameshi, O. B, Supeno, Astutik, S 2020 Pengaruh model pembelajaran inkui terbimbing terhadap kemampuan komunikasi ilmiah dan hasil belajar fisika siswa SMA Jurnal Ilmu Fisika dan Pembelajarannya 4(1) 21-30
[17] Supeno, S, Kurnianingrum, A M, & Cahyani, M U 2017 Kemampuan penalaran berbasis bukti dalam pembelajaran fisika Jurnal Pembelajaran Dan Pendidikan Sains 2(1), 65–78
[18] Vijayanto, T, Supeno, Bektiarso, S 2020 Pengaruh model inkui terstruktur terhadap kemampuan scientific explanation siswa dalam pembelajaran fisika di SMA Jurnal Pendidikan Fisika Tadulako Online 8(2), 18-24
[19] Falk, B and Ort, S 1998 Sitting down to score: teacher learning through assessment The Phi Delta Kappan 80(1) 59-64
[20] Gambell T and Hunter D 2004 Teacher scoring of large-scale assessment: professional development or debilitation? Journal of Curriculum Studies, 36(6), 697-724
[21] Gearhart, M and Osmundson, E 2009 Assessment portfolios as opportunities for teacher learning. Educational Assessment 14(1) 1-24
[22] Carlson, D, Borman, G, & Robinson, M 2011 A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement Educational Evaluation and Policy Analysis 33(3) 378–398
[23] McNaughton, S, Lai,M, & Hsiao, S 2012 Testing the effectiveness of an intervention model based on data use: a replication series across clusters of schools. School Effectiveness and School Improvement 23(2) 203–228
[24] Mandinach, E B 2012 A perfect time for data use: using data-driven decision making to inform practice. Educational Psychologist 47(2) 71-85
[25] Athanases, S Z, Wahleithner, J M, & Bennett, L H 2012 Learning to attend to culturally and linguistically diverse learners through teacher inquiry in teacher education. Teachers College Record 114(7) 50
[26] DeLuca, C and Bellara, A 2013 The current state of assessment education aligning policy, standards, and teacher education curriculum. Journal of Teacher Education 64(4) 356-372
[27] Marsh, J A 2012 Interventions promoting educators’ use of data: research insights and gaps. Teachers College Record 114(11) 1-48
[28] Wayman, J C, & Jimerson, J B 2013 Teacher needs for data-related professional learning. Studies in Educational Evaluation 42 25-34
[29] Frank, M., Walker, J, Attard, J, & Tygel, A 2016 Data literacy: what is it and how can we make it happen? The Journal of Community Informatics 12(3) 4-8
[30] Winkelmann, K, Baloga, M, Marcinkowski, T, Giannoulis, C, Anquandah, G, and Cohen, P 2015 Improving students’ inquiry skills and self-efficacy through research-inspired modules in the general chemistry laboratory. J. Chem. Educ. 92(2) 247–255