LESSONS AIMED AT DEMONSTRATING STATISTICAL LITERACY SKILLS: A CASE STUDY OF JAPANESE HIGH SCHOOL LESSONS ON READING STATISTICAL REPORTS

Shunya Koga
Kwassui Women's University, Faculty of Wellness Studies, Japan
s.koga@kwassui.ac.jp

This study examined whether statistical literacy skills could be demonstrated through high school lessons. First, a statistical literacy process, based on critical thinking, was developed for the interpretation and evaluation of statistical reports. A worksheet, which reflected this process, was provided to help the students read statistical reports. A total of seven 50-minute lessons were conducted with 34 high school students, mainly using this worksheet. In addition, statistical literacy tests were administered at the beginning and end of the lessons to determine whether students' skills improved through the lessons. The results indicated the presence of multiple skills that students demonstrated through the lessons.

INTRODUCTION

Many organizations have produced numerous statistical reports that include the following components: the population studied, method of data collection, and results of analysis. These reports must be critically evaluated to facilitate decision making. In addition, such information is usually accessible through the media. The ability to evaluate such reports is referred to as statistical literacy (e.g., Gal, 2002). To enable students to demonstrate statistical literacy in the future, schooling has attempted to teach statistical literacy in high school. Statistical literacy education has mostly focused on understanding and critically evaluating statistical information (Budgett & Rose, 2017) or assessing statistical literacy (Budgett & Pfannkuch, 2010). This study focused on the method that teaches how to interpret and evaluate statistical reports, namely, what procedures are used to interpret reports and what points to focus on to evaluate the contents.

This study developed a method for students to read and evaluate statistical reports using critical thinking, which is closely related to statistical literacy. Statistical literacy includes elements related to critical thinking skills, such as understanding arguments or identifying claims and evidence. In other words, critical thinking works as a foundation for statistical literacy. Weiland (2017) stated that the use of "critical" in statistical literacy refers to the use of critical thinking. Sharma (2017) reviewed the concept of statistical literacy and stated that "it is evident that statistical literacy is a complex construct that requires not only a range of basic skills (reading, comprehension, and communication) but also higher-order cognitive skills of interpretation, prediction and critical thinking" (p. 129). Ziegler and Garfield (2018) also stated that "definitions of statistical literacy range from the context of basic skills (Garfield et al., 2002) to critical thinking (Gal, 2002)" (p. 161). Statistical literacy includes knowledge elements such as skills and dispositional elements such as attitudes. This study formulated a method to demonstrate statistical literacy skills based on critical thinking. In addition, by demonstrating a continuum of various statistical literacy skills while reading the statistical reports, students are expected to evaluate them appropriately. Thus, the study examined whether the statistical literacy skills based on critical thinking could be demonstrated through the lessons.

DEVELOPMENT OF THE STATISTICAL LITERACY PROCESS

Koga (2022) determined eight statistical literacy skills based on critical thinking skills (skills A through H in Figure 1). Skill development was based on critical thinking skills proposed by Ennis (1996), Facione (1990), and others and are demonstrated while interpreting and evaluating reports. Because critical thinking can be viewed as a series of cognitive processes (Dwyer et al., 2014), using these skills while reading statistical reports may help us to evaluate the reports appropriately. Therefore, a statistical literacy process (Figure 1) was developed including the eight skills arranged accordingly.

To evaluate reports, it is important to understand its contents properly. Statistical reports that are transmitted by the media often contain claims. It is important to consider the situation under which the claim arises to understand its context. To understand texts correctly, it is essential to interpret the meaning of statistical terms (e.g., average and percentage) and ambiguous expressions (e.g., "concerns"

In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11 2022), Rosario, Argentina.* International Association for Statistical Education. iase-web.org

and "relates"). This constitutes the first step of the process. At this stage, skill D comes into play. After this, one must identify the claims that are made in the reports and the evidence that supports them (e.g., statistical results to support a claim). Then, the claims must be evaluated to determine whether they are reliable. Therefore, it is necessary to implement skill A, clarify the claims and evidence, and assess the reliability of the evidence using B.

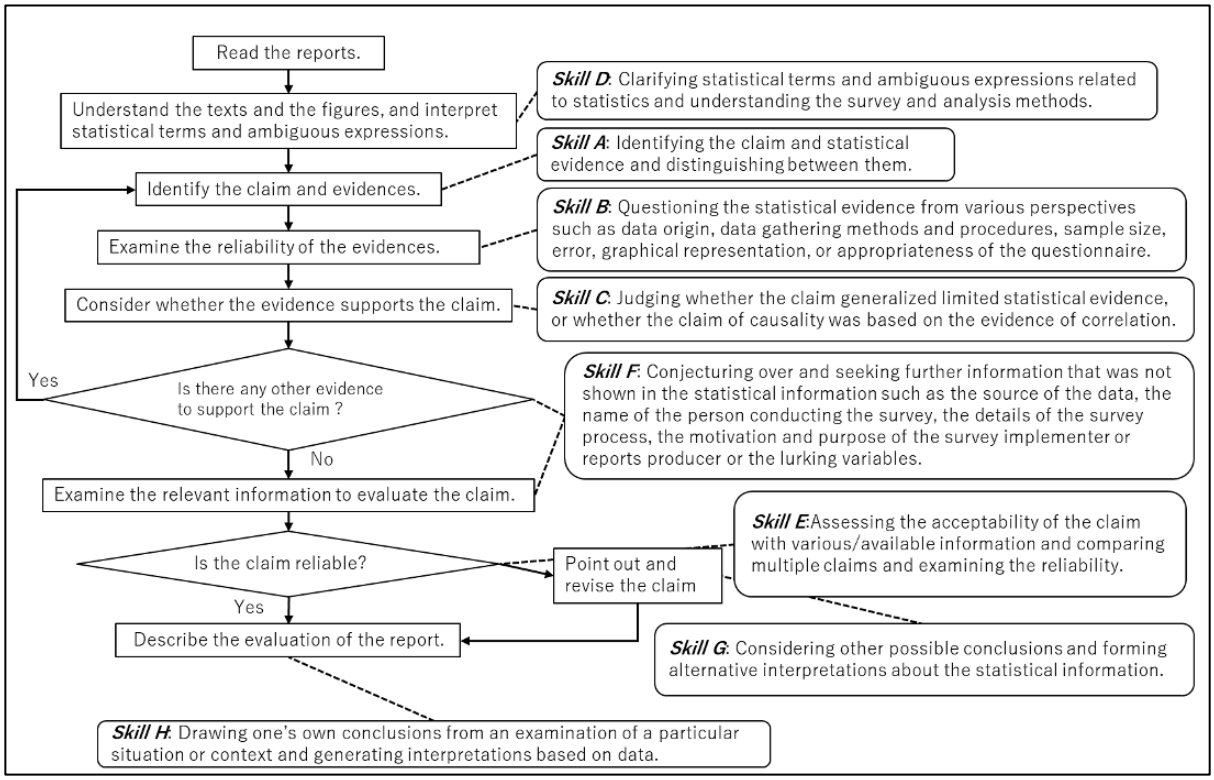

Figure 1. Statistical literacy process

A claim cannot be considered reliable simply because its evidence appears reliable; the strength of the evidence must also be determined. It is necessary to examine whether the evidence is sufficient to support the claim. Therefore, it is essential to assess how the statistical evidence relates to the claim. Thus, the claim's reliability can be judged by the evidence's reliability and relevance. This requires skill C. In addition, skill F is performed to determine whether the claim is reliable by identifying additional evidence from the reports or relevant information from other sources. Subsequently, learners implement skills E and H to state whether the claim is reliable, along with adequate and logical justification for evaluation. If the claim is unreliable, then skill G suggests how the research conducted in the report can be improved, namely how the sampling method and research design be improved. Dwyer et al.'s (2014) framework suggests that monitoring the elements involved in critical thinking is an important aspect of critical thinking. For example, critical thinkers should summarize their thoughts and assess whether they are coherent. Meanwhile, critical thinkers should reflect on how they evaluated the information and identify problematic aspects to identify areas for improvement.

LESSON DESIGN

This study focused on lessons wherein students evaluated statistical reports according to a worksheet that reflected the process (Figure 1). The outline of the lessons (seven lessons in total) is shown in Table 1 (each lesson was 50 minutes). All reports used in the lessons were based on those submitted to *Evaluate Statistically Based Reports* (ESBR) external testing in New Zealand. The Year 13 (ages 17–18) unit on ESBR includes discussions of survey design, margin of error, and possible sources of bias (New Zealand Qualifications Authority, 2019). Further, ESBR is based on non-sampling errors and causal relationships. The author selected statistical reports from the external tests that were

appropriate for the content of each lesson in this study. Then, the reports were translated into Japanese and slightly adapted to the Japanese contexts.

Table 1. Outline of the lessons

Lesson	Reports Shown to Students	Student Activities	Statistics-related Content
1		Answer the pre-test. Learn basic knowledge about survey methods. Consider how to select a sample.	• Target population • Sample • Representative samples
2	Article on investigations of fines and illegal parking	Consider whether the location of the parking spaces investigated is appropriate from the perspective of sampling.	
3	Article on investigations of illegal drivers by a life insurance company	Consider the validity of a sampling frame that investigated the customers of an insurance company. Examine the validity of the accuracy of the responses, focusing on the double-barrel and ambiguous wording of the questions.	• Double-barrel and ambiguous wording of questions
4 & 5	Article on investigations of the relationship between students' part-time hours and grades	Examine the validity of the accuracy for the responses, focusing on the double-barrel and ambiguous wording of the questions. Practice data analysis in spreadsheets using the methods described in the article. Learn how to read a scatterplot and correlations. Learn about the difference between correlation and causation from the perspective of temporal order and confounding variables.	• Scatterplot • Correlation • Causation • Explanatory and Response variables • Confounding variables
6 & 7		Search the report and evaluate it. Answer the post-test.	

This study involved one class with 37 third-grade students (17–18-year-olds) from a private high school. They were enrolled in a regular course in Japan. In addition, most students go on to higher education after graduation. In Japan, school-based statistics education generally involves learning data analysis and the use of software to calculate statistics. Until recently, the Japanese curriculum did not focus on statistical literacy. Accordingly, most Japanese high school students (15–18-year-olds) have not received sufficient statistical literacy education. This was the first time that the students learned statistical literacy skills, such as evaluating statistical reports. The lessons were conducted by the teacher who had been instructed by the author. A total of seven 50-minute lessons were delivered from June 2021 to July 2021.

A statistical literacy test was administered to examine whether students demonstrated statistical literacy skills. This test was comprised of two sections (Figure 2). Both sections contained statistical reports, aimed at measuring skills B, C, E, and F. Students completed the same test twice: before Lesson 1 (pre-test) and after Lesson 7 (post-test). On this test, students were required to read the reports and judge the quality of report contents on a four-point scale ranging from "good" to "not good." After that, they had to state in detail the reasons for their rating. Students who demonstrate the skills shown in Figure 1 should provide descriptions similar to those displayed in Table 2. Thus, the students' pre-test and post-test descriptions were examined by the author and teacher to verify whether students were able to demonstrate each skill listed in Table 2 through the lessons. If a student's description was similar to the description for each skill listed in Table 2, it was assumed that the skill was demonstrated.

Read the following report and choose one option from "good to not good" as to how good this article is. In addition, enter the reason for your decision in the Google Forms.

For young people, time spent on SNS affects physical activities

Dr. Q, a professor at the University of P, conducted a survey of 150 students at universities in Tokyo in mid-August to show that the amount of time young people spend on social networking services (SNS) has an impact on their exercise time. Dr. Q asked them "How much time did you spend on social networking sites this week?" He also asked the same question about physical activities time and collected the responses.

The result of the survey is shown in the figure below. Based on the result, Dr. Q said, "The more time spent on SNS, the less time that is spent on physical activities. In other words, the time young people spend on SNS affects the time they spend physical activities." Furthermore, he said, "For an hour spent on SNS, physical activities time will decrease by about 30 minutes."

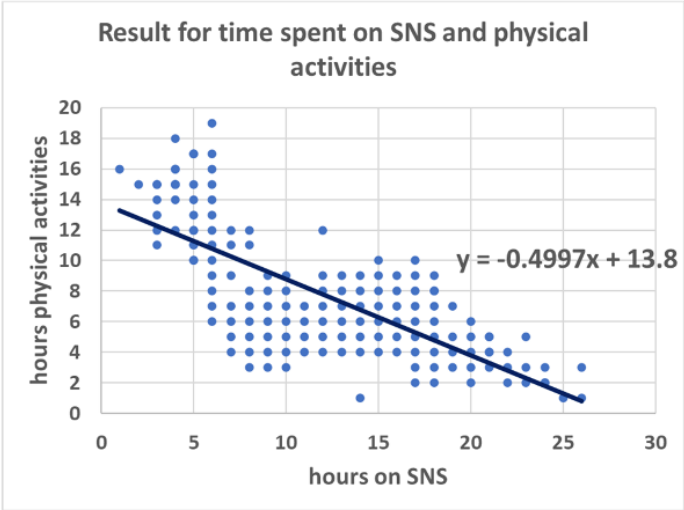

Figure 2. Task from the statistical literacy skills test

RESULTS

Data from 34 students who participated in all lessons and worked on all tasks were included in the analysis. In the pre-test before Lesson 1, approximately 40% of all students (13/34) answered "good" or "somewhat good" to the question at the beginning of the task. When the teacher showed these results to the students, the author could confirm that they were wondering why their evaluations were so different when they were reading the same article. The teacher then explained the importance of properly evaluating statistical reports, which motivated the students. Consequently, from Lesson 2 to Lesson 5, the students were able to complete the worksheets correctly with the teacher's help as needed. It was thought that through these lessons, the students became aware of the flow of the process in Figure 1. In Lesson 6 and Lesson 7, the students were required to searched for reports they were interested in or reports on current events. They were encouraged to interpret the content from different perspectives. After Lesson 7, most students answered "somewhat not good" or "not good" on the post-test. This indicated that the students expressed a more critical view of the reports or believed that the content of the reports was incorrect.

Results from the pre-test and post-test are displayed in Table 3. The author and teacher counted whether each skill appeared in students' descriptions. After counting, the McNamar test was conducted to determine whether there was a change between pre- and post-tests. The results showed that some skills changed between pre- and post-tests whereas others did not. Skills B, C1, and E were changed.

During the lessons, the students spent sufficient time on activities such as assessing the reliability of evidence and exploring the connections between claims and evidence using worksheets. Therefore, it is possible that these skills were promoted or that students could better express these skills. However, skill C2, which focused on the difference between causality and correlation, did not change for 23 students. Although the lessons covered the same content, students find it difficult to understand correlation and causality (e.g., Garfield et al., 2008). It is also possible that the students' understanding may not have been completely correct. In addition, there was no change in skill F for both sections. In evaluating reports, the students engaged in suggesting alternatives or revisions to the claim; however, this was not the case in the tests. It can be said that students' perception of the term "evaluation" is limited to judging the content of reports as good or bad.

Table 2. Test evaluation perspectives in section 1

Skill in Figure 1	Specific Skill Targeted (in <>) and Example of Student Responses
B	<Students were able to examine the reliability of the evidence by focusing on the validity of the time of day to ask questions.> "In this survey, respondents were asked about their social networking and exercise time for the week on weekends, which may not be reliable for the time of response data because the respondents' memories may be fuzzy."
C1	<Students were able to consider whether the sample was representative of the population indicated in the claim.> "Although this survey only included university students in Tokyo, it is unlikely that the results of this survey can be adapted to include all young people, since the claim targets 'young people.'
C2	<Students could check to see if they were not making a causal claim based on the evidence of correlation.> "The results of this study can be interpreted as a correlation but cannot be used to make a causal claim."
E	<Students were able to infer information that was not presented in the article and use that information to judge the credibility of the claims and evidence.> "The researcher or author of the article may have a desire to limit the time spent on social networking sites; A third variable would be motivation to learn and parental education."
F	<Students could suggest alternatives to the survey and revisions to the claim that would be appropriate.> "The claim should be changed to, 'There may be a correlation between social networking and exercise time among college students in Tokyo.'"

Table 3. Pre- and post-test results

Skill *B*		Post	
		No D	D
Pre	No D	19	13
	D	2	0

$p = 0.0074$

Skill *C1*		Post	
		No D	D
Pre	No D	6	21
	D	2	5

$p = 0.0001$

Skill *C2*		Post	
		No D	D
Pre	No D	23	9
	D	1	1

$p = 0.0215$

Skill *E*		Post	
		No D	D
Pre	No D	9	19
	D	2	4

$p = 0.0002$

Skill *F*		Post	
		No D	D
Pre	No D	26	2
	D	5	1

$p = 0.4531$

"D" means "Description" and "No D" means "No Description"

FUTURE CHALLENGES

This study showed that the students were able to demonstrate some statistical literacy skills based on critical thinking while evaluating statistical reports using the statistical literacy process. However, this was a quasi-experimental study, wherein the students answered the same test twice to serve as their own control group. Therefore, it is possible that the post-test was easier for students to answer as they became familiar with the questions. In other words, it is difficult to determine the lessons contributed fully to the students' demonstration of their statistical literacy skills. Additional techniques, such as interviewing students who participated in the lessons, must be utilized in future research.

Further, additional lessons should be designed to address statistical literacy. For example, activities on how to conduct a survey and communicate the results of reports should be incorporated into the lessons of this study. Writing a report is likely to enhance knowledge of statistical terminology and descriptions of statistical evidence and claims. Weiland (2017) noted that "being literate in the reading context can only go so far, as reading often operates in dialog with writing, and some experience in writing is important to be able to make sense of and evaluate statistical arguments" (p. 38). The students who participated in this study had no experience with conducting a survey. In the future, students who have previous experience with such activities and who have learned how to answer questionnaires and analyze reports may better understand the content of these lessons.

ACKNOWLEDGMENT

All procedures were approved by the ethics committee of the Faculty of Human Sciences, University of Tsukuba (No: 2021-25 A). In addition, this work was supported by JSPS KAKENHI (Grant-in-Aid for JSPS Fellows) Grant Number JP19J20055. I would like to thank Dr. Naohiro Higuchi from the University of Tsukuba for valuable discussions. Furthermore, I would like to thank all the students who participated in the lessons of this study.

REFERENCES

Budgett, S., & Pfannkuch, M. (2010) Assessing students' statistical literacy. In P. Bidgood, N. Hunt, & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 103–121). John Wiley and Sons. https://doi.org/10.1002/9780470710470.ch9

Budgett, S., & Rose, D. (2017). Developing statistical literacy in the final school year. *Statistics Education Research Journal*, *16(1)*, 139–162. https://doi.org/10.52041/serj.v16i1.221

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43–52. https://doi.org/10.1016/j.tsc.2013.12.004

Ennis, R. H. (1996). *Critical thinking*. Prentice Hall.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction—research findings and recommendations*. American Philosophical Association.

Gal, I. (2002). Adult's statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–25. https://doi.org/10.2307/1403713

Garfield, J., Ben-Zvi, D., Chance, B., Medina, E., Roseth, C., & Zieffler, A. (2008). *Developing student's statistical reasoning: Connecting research and practice*. Springer. https://doi.org/10.1007/978-1-4020-8383-9_14

Koga, S. (2022). Characteristics of statistical literacy skills from the perspective of critical thinking. *Teaching Statistics*, *44*(2), 59–67. https://doi.org/10.1111/test.12302

New Zealand Qualification Authority. (2019). *Achievement standard (91584 Evaluate statistically based reports*). https://www.nzqa.govt.nz/nqfdocs/ncea-resource/achievements/2019/as91584.pdf

Sharma, S. (2017). Definitions and models of statistical literacy: A literature review. *Open Review of Educational Research*, *4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

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

Ziegler, L., & Garfield, J. (2018). Developing a statistical literacy assessment for the modern introductory statistics course. *Statistics Education Research Journal*, *17*(2), 161–178. https://doi.org/10.52041/serj.v17i2.164