Statistical reasoning or statistical method: Students’ preferences for learning Statistics

Kusumarasdyati

English Department, Universitas Negeri Surabaya, Indonesia
e-mail: kusumarasdyati@unesa.ac.id

Abstract. Statistical reasoning should precede the statistical method in the teaching of statistics because the students should master the concepts before they perform computations. To find out whether the students were aware of this principle, the present study explored their preference and the underlying reasons. Twenty undergraduates were taught central tendency and variability by using statistical method (with machine computation) and statistical reasoning (with manual computation) respectively in one session. Most of them turned out to prefer to reason and compute numeric data manually because they understood the statistical concepts and the process of the computation. On the other hand, the rest of the undergraduates preferred the teaching of statistics using the statistical method with machine computation for three reasons: accuracy, efficiency, and convenience. The implications of these findings for the teaching of statistics were presented.

1. Introduction

The teaching of statistics can be approached in two ways: statistical reasoning and statistical methods. Statistical reasoning means “the way people reason with statistical ideas and make sense of statistical information” [1]. Approached in this way, teaching statistics encourages the students to understand the concepts, link one concept to another, and make correct interpretations about the results of the computation. On the other hand, statistical methods refer to the techniques of crunching a large number of data to make them meaningful. Teachers who focus on statistical methods in the classroom only develop the students’ skills in computing numbers and probably create graphs based on the results of the computation.

It is important to note, however, that there is more to statistics than computation. Statistics is “the science of collecting, organizing and interpreting data” [2] to solve problems in many aspects of life, and the computation is needed in organizing the data. To make the results of the computation meaningful, they must be interpreted carefully. The students definitely should understand the relevant concepts to make accurate interpretations and arrive at sensible conclusions. Unfortunately, some teachers lack knowledge about this. They erroneously equate statistics with computation only, teaching the procedures and the computation without training the students to reason [3]. Moreover, in this age of computerization, some teachers tend to teach the students to operate statistical software rather than facilitating them to understand statistical concepts and reason [4, 5].

Bradstreet [6] argued that for better learning outcomes, statistical reasoning should be taught first, and statistical methods—usually utilize the software for more convenient computation—follow after the students have mastered the statistical concepts. This principle triggers important questions: Do the students know that they should acquire the concepts well before learning to compute numeric data? Do
they think they learn statistics better by reasoning the concepts, or are they content with using software to perform computations? The answers to these questions are vital to ensure effective learning of statistics. If the students are aware of the importance of statistical reasoning, the learning objectives are met once they can reason. On the other hand, if they think that learning statistics is completed upon having the ability to run statistical software, the teacher should take necessary actions to make them realize that computations using the software means to an end rather than an end itself.

The present study aimed to explore how statistics should be learned from the students’ perspective. The research questions could be stated as follows: (1) Which did the students prefer in learning statistics: statistical reasoning or statistical method? (2) What was the underlying reason for their preference? The procedures for conducting the study to answer these questions will be explained in the next section. Then the results of the study will be reported and discussed. The last part of this paper will present the conclusions drawn from the study and present the implications for the teaching of statistics.

2. Methodology
This study was descriptive research which was approached qualitatively. The participants consisted of twenty undergraduates majoring in the teaching of English as a foreign language (EFL). They took the course “Educational Research Methodology” at the English Department in 2018. In the following semester, they would take the course "Quantitative and Qualitative Analysis" where they were going to learn statistics, so this study attempted to find out whether they knew that there was more to statistics than mere computation.

The data were collected from the students in one session lasting for 100 minutes. This session was divided into two parts: the statistical method (30 minutes) and statistical reasoning (70 minutes). In the first part, the students were taught how to compute central tendency (mean) and variability (standard deviation) using excel. Each student ran this software on his or her laptop and followed my instructions about the procedures of computing the mean and the standard deviation. Afterwards, I taught them how to interpret the results of machine computation. After the first part of the lesson ended, the second part immediately followed. In the second part, the topics remained the same but the way that the materials were delivered entirely differed. The lesson began with explanations about what central tendency and variability meant, what their functions were, and why the students needed to know about them. After that, I showed them how to compute the mean and the standard deviations on the whiteboard step by step, explaining the concepts that came up during the manual computation. When the results of the computation were obtained, I demonstrated how to interpret them by drawing a normal curve with the symbols of mean and standard deviation in the appropriate places and giving detailed accounts of what they meant.

As soon as the lesson ended, the students completed an unstructured questionnaire, in which they had to state whether they preferred the first method of learning statistics (machine computation) or the second one (manual computation) when they learned the two topics for the first time. They also had to elaborate on all of the reasons for their preference. The data were analyzed by following these procedures: familiarizing & categorizing, coding & reducing, interpreting & representing [7].

3. Results and Discussion
The results showed that most of the students preferred manual computation when they learned concepts for the first time, indicating they acknowledged the importance of statistical reasoning. Only a few of them supported statistical method, in which the computation became the emphasis of the instruction rather than understanding the concepts. The reasons underlying their preference varied and would be reported below. Because they were still learning English as a foreign language, their answers contained some grammatical errors. These errors, however, did not seriously affect the meaning of the whole sentences.

The reasons for preferring statistical method using software were threefold: accuracy, efficiency, and convenience. The students claimed that the results of the machine computation were more accurate, eliminating the chance of errors. One of the students, Fred, wrote that, “...the result is more accurate
than the manual method” and Lisa detailed it further by writing “I think it is easier to compute data using software because it avoided us from human error like when we forget the formula, wrong in counting the scores and typo in computing the scores.” It was obvious that machine computation ensured the accuracy of the computation results. The students did not need to worry about making mistakes due to the failure in memorizing complex formulas, miscalculating the numeric data and mistyping the numbers. The second reason was efficiency, as Lisa put it, “… the use of software can help us to compute the data and get the result faster than compute the data manually.” Similarly, George stated, “… it is not time-consuming …” and Chloe wrote, "… it can save time to compute the scores." For these students, it was better to use the software when learning to compute the mean and the standard deviation for the first time because it was more efficient concerning the time spent on the computation. After entering the scores in the computer program, they could obtain the results within a few seconds by clicking certain buttons. The last reason was closely related to efficiency, i.e. convenience. George, for instance, described how convenient it was to learn statistics using the software as follows: "I think using software for calculating the average and standard deviation is much easier. Because we only need to compute/type the formula, and then we can get the result.” Likewise, Fred wrote that “In my opinion for the first computing standard deviation it easier use the computer with the Microsoft Excel formula.” Finally, Chloe thought that "We do not need a quite complicated formula. In the software, the formula already exists.” According to them, using software to learn statistical computation not only saved a lot of time but also provided convenience. Machine computation was easier for them because they did not need to write the formulas and go through the complicated process of calculating the numbers.

On the other hand, the students who preferred to learn the central tendency and variability by manual computation gave two reasons: concepts and process. It was mentioned earlier in this paper that understanding concepts were central in learning statistics and apparently, most students who chose the manual computation over the machine one were aware of this, as shown in the reasons they gave below.

Lucy : For the first time I learn about how to conduct a standard deviation for qualitative research, it is better to learn manually because I will know the concept in depth.
Kent : I think it is easier to study manually for the first time both for means and standard deviation. Because I need to know the concept first and the formula.
Jane : In my opinion, the manual learning material is quite better than using a computer. In manual statistic, we can learn about the background of what is exactly "mean" and "standard deviation" through the formula given by the teacher.

Lucy, Kent and some other students stated explicitly that knowing the concepts was essential when learning statistics. Jane even explained that she understood the concepts of mean and standard deviation by learning the use of the formulas in manual computation. By using the formulas, they got a full grasp of not only the concepts but also the computation process. They revealed this as follows.

Mary : For me, to understand the mean and standard deviation, it is easier to compute the formulas manually. Because it gives me the detail of formulas step by step, while if we use the machine, we only get the result without knowing the detail about how we can get the number. So, it’s better to use the formulas manually for a beginner.
Rick : For the first computing the standard deviation and mean, I think I prefer to use manual way because even though that’s more complicated than using the software I can know clearly where the number from. I can understand what are the components that I need to get the standard deviation and the mean.
Prue : Because of already learning how to compute the mean of data in high school, I think it is easiest to compute mean by using the software of a computer. However, for learning about standard deviation that unfamiliar, I choose to compute manually. It is because I am going to more comprehend the concepts of standard deviation, the steps to compute it, what it is
about, what the function of it. If I do not know the concepts of it and only depend on software, I will find many difficulties in conducting my future research.

Lena: Easier to understand when learning: manual computation. Because I can know the original formula and I can know the step by step such as where those numbers come from comparing to by using software computation. Software computation is easier to use, but I feel like I am “blind” when I use it. It is simple but I don’t know where are those numbers come from. I just automatically know the result.

Obviously, for the students, it was compulsory to know how they could obtain the numbers showing the mean or the standard deviation of a distribution of scores, and the only way to know it was by performing manual computations. The process of the computation assisted them to find out where the numbers came from. For example, by computing the mean they learned the terms the sum of scores and the number of scores, how to use them in the formula, what they obtained as a result of the computation, and how to interpret the result. This process helped them to accumulate pieces of information gradually to construct knowledge about the mean, achieving a good understanding of this concept in the end. The ease of use that the software offered did not deter Rick and Lena from preferring the manual computation. Lena even considered it puzzling (or ‘blind’, as she called it) to get the result instantly from Excel\textsuperscript{TM} without knowing the process of computation.

It was interesting that the students who preferred machine computation and those who liked manual computation better gave very different reasons. The choice of the former students was motivated by practicality, i.e. the convenience in use and the instant results. It means they viewed learning statistics as gaining computation skills only, indicating that they supported the statistical method. The latter group of students, in contrast, concerned themselves more with conceptual understanding. For them, it was vital to reason and gain knowledge about the statistical concepts, and manual computation became a useful tool to promote their understanding of the concepts. It could be inferred that they put more emphasis on statistical reasoning.

Traditionally, statistics was taught using a statistical method, which enabled the students to compute numeric data. In 1990, however, the teaching of statistics began to undergo reformation and the prominence shifted from computation to conceptual understanding [8]. Therefore, the students who expressed their preference for manual computation to know the concepts and the process of calculating numbers were on the right track in learning statistics. The manual computation was indeed an effective tool to achieve conceptual understanding because it developed their statistical intuitions about the best technique to use, their self-reliance in solving the problems and their quantitative skills [5]. It explained why Mary stated that “if we use the machine, we only get the result without knowing the detail about how we can get the number” and Lena argued, “Software computation is easier to use, but I feel like I am ‘blind’ when I use it. It is simple but I don’t know where are those numbers come from. I just automatically know the result.” They felt they were deprived of the opportunity to practice their quantitative skills and sharpen their statistical intuitions. For them, manual computation made learning statistics more meaningful and effective.

It does not necessarily mean that machine computation has no place in the teaching and learning of statistics. In fact, the students should be able to operate Excel\textsuperscript{TM} or other similar computer programs to get accurate results within a very short time when they conduct research in the future, but they may do it only after they have mastered the statistical concepts by learning how to compute manually. It is important to underline that statistical reasoning precedes statistical method [6], so manual computation must be learned before machine computation to ensure conceptual understanding.

4. Conclusion
This study attempted to explore the undergraduates' preference in learning statistics, i.e. statistical reasoning by utilizing manual computation or statistical method by using machine computation. Most of them preferred to compute manually because they could learn the concepts and going through the process of the computation could enhance the conceptual understanding. The rest, however, preferred
to compute the numeric data by utilizing software because the results of the computation were more accurate, could be obtained more rapidly, and were easier to get without any complicated process of calculating numbers.

These findings have practical implications for the teaching of statistics. Teachers need to make the students aware of the two approaches—statistical reasoning and statistical methods—in the introductory session of the statistics class. Further, teachers should explicitly emphasize the importance of statistical reasoning when students learn new concepts in statistics. After communicating these verbally to the students, the teachers could begin the instructions by applying statistical reasoning in facilitating the students to get a full grasp of the concepts, using only manual computation to introduce the concepts. When the mastery of concepts and manual computation has been achieved, the teachers may introduce various types of statistical software to ease computation.

References
[1]  Garfield J and Ben-Zvi D 2007 How students learn Statistics revisited: A current review of research on teaching and learning Statistics International Statistical Review 75(3) 372–396
[2]  Bennett J, Briggs W L, and Triola M F 2014 Statistical reasoning for everyday life (Harlow: Pearson)
[3]  Ben-Zvi D and Garfield J 2004 Statistical literacy, reasoning, and thinking: Goals, definitions, and challenges The Challenge of Developing Statistical Literacy, Reasoning and Thinking (Dordrecht: Kluwer Academic Publishing)
[4]  Pimenta R 2006 Assessing Statistical reasoning through project work Int. Conf. on Teaching Statistics (Salvador)
[5]  Khamis H J 1991 Manual Computations—A tool for reinforcing concepts and techniques The American Statistician 45(4) 294-299
[6]  Bradstreet T E 1993 Teaching Introductory Statistics Courses so that nonstatisticians experience statistical reasoning The American Statistician 50(1) 69-78
[7]  Ary D, Jacobs L C, and Sorensen C K 2010 Introduction to Research in Education (Belmont: Wadsworth, Cengage Learning)
[8]  Tu W and Snyder M M 2016 Developing conceptual understanding in a Statistics Course: Merrill’s first principles and real data at work Educational Technology Research and Development 65(3) 579-595