Statistical Reasoning Skills and Attitude: The Effect of Worked Examples

Li Jin, Yoon Jeon Kim, Marilyn McGhee, and Robert Reiser
The Florida State University

Abstract

The purpose of this study was to examine the effects of worked examples on learners’ statistical reasoning skills and their confidence and satisfaction levels. Twenty-six graduate students were randomly assigned to two groups. After being provided with the same instructional material on the general rules of comparing group means and how to make heuristic inferences based on data without computing a t-test, one group studied worked examples and the other group practiced problem solving with no worked examples. A posttest on statistical reasoning skills executed after the instruction revealed that participants who studied worked examples outperformed the problem practice group (p<.05). The survey on participants’ confidence and satisfaction level indicated that participants who studied worked examples were more confident than their counterparts (p<.05) but showed no difference on satisfaction level. These results are discussed in terms of how the use of worked examples changed the process of learning of statistical reasoning skills. The implications of the use of worked examples in statistics education are discussed as well.

Keywords: Statistics education, worked example, statistical reasoning skill, teaching statistical reasoning

Statistical Reasoning Skills and Attitude: The Effect of Worked Examples

Statistics education has drawn increasing attention in recent years. First, tremendous numbers of students have enrolled in statistics course as a requirement for their degree programs (Garfield & Ben-Zvi, 2005). This is because statistical knowledge has become requisite for a wide range of fields of study (Garfield & Ahlgren, 1988). Second, there is a strong call for reform of statistics instructions. Most instructors tend to teach concepts and procedures rather than statistical reasoning. As a result, students don’t know how to apply statistical concepts and rules to concrete problems or make inferences even when they can earn good grades on statistics courses (Garfield, 2002). Third, there are broad concerns regarding students’ attitude towards statistics courses. Statistical courses are viewed by many students as the most difficult in their curricula, which are regarded as obstacles standing in the way of attaining their degrees (Gal & Ginsburg, 1994).

In recent years, more and more educators have advocated that the focus of statistical education should be developing statistical reasoning skills rather than letting students memorize definitions, rules, and procedures (Chance, Garfield, & delMas, 2000). Statistical reasoning is the way people reason with statistical ideas and make sense of statistical information (Garfield, 2002; Garfield et al., 2005). Statistical reasoning may involve connecting one concept to another, such as means and spread, or combining ideas about data, such as frequency and chance. Statistical reasoning also means understanding and being able to explain statistical processes as a whole, and being able to interpret statistical results, such as explaining the meaning of confidence intervals and significance tests. Statistical reasoning skill is one of mostly desired outcome of a statistics course (Garfield, 2002).

Teaching statistical reasoning is challenging because little is known about the statistical reasoning process or how students actually learn statistical ideas (Hawkins, Jolliffe, & Glickman, 1992). The knowledge of how to design and implement effective instructional strategies to teach statistical reasoning skills is of importance. However, the majority of the studies conducted on teaching statistical reasoning have focused on identifying students’ misconceptions and the dynamics of the teaching-learning process. Only a few studies have addressed the aspect of the instructional activities and methods that are effective in statistical education.

Although only a few studies exist, well-designed instructional activities and methods have been shown to be effective in facilitating teaching and learning in statistical education. For example, Schwarz and Sutherland (1997) and delMas, Garfield, and Chance (1999) reported benefits of using simulations to improve students’ statistical reasoning. According to some other authors (Giraud, 1997; Keeler & Steinhorst, 1995; Magel, 1998), small group activities appeared to work better for statistics learners when they worked together with 2 or 3 peer students to solve problems, discuss procedures, and work on projects. Bradstreet (1996) suggested using real data and questions related to students’ experience for instruction because students would understand the methods as they use them to
solve real world problems instead of memorizing a list of isolated formulas and terms. Bradstreet (1996) also suggested using graphics to communicate statistical concepts, procedures, and results. He argued that graphics can enhance understanding for intro-level statistics learner.

While there are many good instructional activities and methods that have been tested as effective in many disciplines, they will not necessarily assist students to correctly understand and reason about statistical concepts (Chance, 2004). There is a strong need for empirical studies on the effectiveness of utilizing well-established instructional strategies in the field of teaching statistical reasoning.

Learning from worked examples is one of the instructional strategies that have been reported being effective in many disciplines (Cooper & Sweller, 1987; Darabi, Nelson, & Palanki, 2007; Darabi, Sikorski, Nelson, & Palanki, 2006; Paas & Merrienboer 1994; Sweller & Cooper, 1985; van Gerven, Paas, van Merrienboer, Hendriks, & Schmidt, 2003). Worked examples provide an expert’s problem solution for a learner to study. They typically include a problem statement and a procedure for solving the problem (Atkinson, Derry, Renkl, & Wortham, 2000). According to cognitive load theory research, worked examples, compared to conventional problem-solving approach, can reduce the cognitive load imposed on learners. In contrast to the conventional notion of the best way to teach problem-solving being to give students a lot of problem to solve, Sweller and Cooper (1985) argue that learning from worked examples is a more effective way for novice learners to learn problem-solving. When novice learners are presented with problems, they tend to use novice strategies, such as trial and error, which are very demanding on cognitive resources. Students learn from worked examples, however, often employ more efficient problem-solving strategies they learn from the examples and focus on content of the problems, which makes better use of the cognitive resources. Many studies have reported results in favor of worked example instruction rather than problem-solving practice (Carroll, 1994; Ward & Sweller, 1990; Zhu & Simon, 1987).

Considering the nature of worked examples, we think they are good instructional tools for learning statistical reasoning skills. Studies show that students tend to respond to statistical problems by substituting quantities into computational formulas or procedures without forming an internal representation of the problem (Garfield & Ahlgren, 1988) when they work on problem solving practices. This problem can be solved by presenting students with worked examples. When presented with worked examples that demonstrate how an expert makes sense of statistical data and results, students should be able to learn the rationale for the procedure and how concepts can be applied in new situations.

One other difficulty in learning statistical reasoning skills is that students seem to rely on incorrect intuitions and misconceptions to make an inference or judgment no matter how successfully correct conceptions have been taught (Tempelaar, Gijselaers, & van DerLoeff, 2006) and those fallacies are very stubborn (Konolde, 1989; Lecoutre, 1992). One reason why this difficulty is hard to overcome is that students seldom reflect on the procedures they go through in making inferences and compare their procedures to those of experts. By studying worked examples, students can learn experts’ skills of using correct conceptions and statistical rules in solving problems and compare those skills with theirs. We think that worked examples would help students learn experts’ statistical reasoning processes and check for errors in their own thoughts. This is likely an efficient way for learning complex concepts and skills.

At the same time, one other important thing in helping students become successful in learning statistics is to get them to feel confident about their ability to learn the content and like the content. We think that the use of worked example in statistical education would help solve the problem of negative attitude of many of the students. Some studies report that students felt more confident in learning when they learned from worked examples (Cripper & Earl, 2005; Miller, 2009). Researchers also report that students were better engaged in a worked-example learning situation than they were in a conventional problem-solving process (Darabi et al, 2006). Also, students do not like statistics mostly because they feel it is hard to learn (Gal et al., 1994). Worked examples are believed to be able to reduce cognitive load imposed on learners (make the learning easier) than the traditional problem solving approach (Sweller & Cooper, 1995). When students feel less difficult in learning they may be able to feel more confident and be willing to learn.

In spite of the broad adoption of worked examples in other disciplines, no study has been done on the use of worked examples in statistics education. In this study, we investigated the impact of worked-examples on students’ learning of statistical reasoning and their attitude toward instruction. Specifically, we compared students’
performance and attitude after they learned from worked examples or worked on problem-solving practice. The worked examples provided an expert’s procedure of interpreting statistical data and test results and making references. For problem-solving practice, participants were required to make their own interpretations and references based on data and test results provided to them. Although correct answers were provided to the practice group, no details on how the answers were reached were provided.

Performance was measured by a written test administered after the instructional activities. The test consisted of two questions which tested the participants’ ability to make correct interpretation and reference based on given data sets and certain statistical test results. Four major points of the reasoning process were examined in each of the two questions. Two aspects of the learner attitude, confidence and satisfaction, were measured by a questionnaire adapted from Keller’s Instructional Materials Motivation Scale (2009). There were nine items focusing on the aspect of confidence and six items examining satisfaction level.

We expected that students who learned from worked examples would score significantly higher on the test and indicate higher confidence and satisfaction level than their counterparts who worked on problem-solving practice. These results were expected because worked examples can reduce the cognitive load of the students and it would be easier for students to learn from worked example than to develop their own ways of reasoning using statistical rules and concepts they just learned. Although correct answers were provided to the problem-solving practice group, they needed to figure out the procedure of reasoning or how they were wrong if they could not come up with the correct answers. Compared to the latter group, the worked example learners did not need to search for appropriate steps for solving the problems and instead they could focus on the content that was important for understanding the problems. Theoretically, less cognitive load was imposed on worked example learners which would have freed the cognitive capacity for learning effort (Sweller, 1988; Sweller & Cooper, 1985). As a result, they would learn better than their counterparts. This result is consistence with many previous studies (Cooper, et al., 1987; Paas, et al., 1994; van Gerven, et al., 2003).

We also expected that students who studied the worked examples would report higher levels of satisfaction and confidence because learning from worked-examples would make the learning easier (less cognitive load) and enhance students’ confidence and performance. As a result of experiencing less difficulty and better performance, students would feel a higher level of satisfaction.

Method

Participants

The participants of this study were graduate students at the Department of Educational Psychology and Learning Systems at a large south-eastern public university. This department was a part of the College of Education. Statistics courses were required for all graduate degree programs in this department. Many of the graduate students in this department started to take statistic courses at the first year in their programs. Most of the students participating in this study took at least one statistics course or was taking a statistics course. Among the collected responses, twenty-six were complete and used for this study. There were thirteen participants in each group.

Materials

The material used for the study presented a heuristic way to understand the meaning of a t-statistic used for mean difference comparisons. There were six points listed in the material. The first point recalled the knowledge of when to use a t-test. The second point described the two parts of a t-statistic—its numerator and denominator. The numerator was connected with between group difference and the denominator was explained by within group difference. The third and fourth points explained the meaning of the numerator and the denominator of the t-statistic in terms of data characteristics (means and standard deviations of the two groups). The fifth and sixth points described the situations of significant and non-significant t-test results and their relations with data characteristics.

Independent variables

The independent variable was the type of instructional activity used to teach students how to conduct statistical reasoning about group difference (t-test) based on data. The first level of the independent variable was to study from worked examples (treatment) and the second level was to use conventional problem practice (control). The participants in the treatment group were provided with two worked examples on how to make inference about group mean differences based on the characteristics of the data without calculating the t-statistics. The participants in the control group were provided with two practice problems and correct answers but with no explanations or steps.
of how the answers were achieved. The participants in the control group were expected to apply the knowledge presented in the study material to solve the two problems and compare their results to the correct answers provided.

**Dependent Variables**

Learning was measured using a posttest. The test was designed to assess the participants’ ability to make reasonable references about group-difference based on data characteristics without calculating a t-test. Specifically, the test presented two scenarios of comparisons of two group means when the means and standard deviations of the two groups were provided. The participants were asked to present their reasoning procedures and conclusions on whether group means were different based on the information given. The participants’ responses to the tests were scored based on the inclusion of important points in the reasoning procedures and the conclusions. One point was given for presentation of each of the following three reasoning procedures: consideration of between group difference, consideration of within group difference, and comparison of between group difference and within group difference. One point was given for getting the correct result. Therefore, a participant got 0–4 points for each problem. The highest possible score for the posttest was eight.

Students’ attitude toward the instructional material was measured by a survey. This survey was adapted from Keller (2009)’s Instructional Materials Motivation Survey. The survey used for this study included fifteen items, six of which were related to learners’ satisfaction and nine of which were related to learners’ confidence. Responses to the items were in the form of 5-point Likert scale. The reliability of the items pertaining satisfaction was 0.80 and reliability of the items pertaining confidence was 0.72.

**Procedures**

The participants were randomly assigned to treatment and control groups after they signed the consent forms. Then they were provided with the instructional material (printed). The participants in the treatment group were provided with two worked examples after the instructional material and the participants in the control group were provided with two problems and right answers to the problems after the instructional material. After that, the participants took the posttest. At last, they were directed to complete the attitude survey which was attached to the posttest. It took each participant about fifteen minutes to finish the task.

**Results**

The outcomes measured in this study were learning and two aspects of learner attitude toward the content: confidence and satisfaction. Learning was measured by a posttest which required participants to conduct statistical reasoning about group means based on the instructional they were provided in the printing material. A review of the data revealed no violation of the assumption of homogeneity variance ($F=.35$ for Levene’s test, $p=.56$). With alpha set at .05 and with 13 participants in each group, the probability of detecting a large difference between means was .63.

The mean score on the posttest for participants who studied worked example was 7.5 with a standard deviation of 1.20. This was significantly higher than the mean posttest score for participants in problem practice group (M=5.3, SD=1.18), $t(24)$ = 4.78, $p<.05$. The effect size is .70, which is large according to Cohen (1988)’s benchmarks. The results supported the hypothesis that the participants would learn better from worked examples than practice on problems.

Confidence aspect of attitude was measured by a 9-item survey adapted from Keller’s Instructional Materials Motivation Survey (2009). The mean confidence score of worked example group (M=4.2, SD=.47) was significantly more positive than that of the problem practice group (M=3.8, SD=.49), $t(24)=2.18$, $p<.05$. The effect size is .41, which is a medium effect based on Cohen’s criteria (1988). The results supported the hypothesis that the participants would have higher confidence level when they learn from worked examples than they practiced on problem solving on their own.

Satisfaction aspect of attitude was measured by a 6-item survey adapted from Keller’s Instructional Materials Motivation Survey (2009). The mean satisfaction score for worked example group was 3.8 (SD = .57) and the mean satisfaction score for problem practice group was 3.6 (SD = .76). Results of a t-test revealed that there was no significant difference in satisfaction scores between the two groups, $t(24) = .64$, $p = .53$. This result did not support the hypothesis that the participants would have more positive attitude toward the content when they were presented with worked examples than when they were provided with practice problems only.
Supporting the primary hypothesis of the study, the mean posttest score of participants who learned from worked examples was significantly higher than the mean score of participants who used a conventional problem-practice approach. This is consistent with the results of some studies on using worked examples in other disciplines, such as algebra (Carroll, 1994), geometry (Paas, et al., 1994), and calculus (Miller, 2009). The worked examples illustrated the step-by-step reasoning procedures for comparing group means. Learners could focus on understanding each of the steps at one time. The comprehension of the logic flow from one step to the next is likely to be much easier when each step is understood. On the contrary, the problem practice group most likely searched for ways to solve the problems based on the statistical concepts and general rules they learned. Especially for novice learners, they needed to figure out the relevant concepts, procedures, and steps to process. Thus, learners in the problem practice group were probably faced with a higher demand on their cognitive resources than the learners in worked example group did. As a result, learners in worked example group grasped the expert way of conducting the reasoning with less learning effort while the learners in the practice group came up with novice solutions (in many cases with errors) at probably much higher cognitive investment.

The hypothesis that predicted a positive effect of worked examples on learners’ confidence level was supported by the data. The participants who studied worked examples expressed a much positive feeling about their capability to learn the content presented in the instructional material than the participants who practiced problem-solving. A possible reason for this finding is that when the steps and logic of solving the problem were clear to the learners, the learners felt confident about their capability to learn the content and solve the problems. But for learners in problem practice group, they were not sure about the solutions they developed. When their results were not the same as the correct answers provided, they might have felt difficulty to figure out what was wrong with their reasoning process and felt less assured about what they learned. As a result, they feel less confident.

The hypothesis predicting a positive effect of worked example on learners’ satisfaction level was not supported by the data. The participants who studied the worked examples did not feel more satisfied than the learners in the problem practice group. This result is probably due to the lack of feedback in the whole process. Satisfaction arises from the positive external consequences of a learner’s behavior or the pleasure of having successfully accomplished a task (Keller, 2009). The learners in both groups in this study did not receive any feedback on how well they performed on the test. Therefore, there was no positive consequence that may lead to boosted feeling of satisfaction. In order to increase learners’ satisfaction level in future applications, we suggest that instructors provide learners with feedback on their performance as soon as possible so as to elicit learners’ satisfaction while retaining their improved confidence.

With the large body of empirical studies on the use of worked examples in many fields of subject, we believe that researchers and instructors who build upon the present study will help provide a richer picture of the benefits worked examples may afford in the field of statistics education. Researchers and instructors may consider using worked examples instead of problem practice not only when they teach statistical reasoning skills but also in situations that they teach other knowledge in statistics courses, such as distinguishing concepts, choosing appropriate tests, and so on. One important suggestion for the use of worked examples is that the instructors should provide feedback to the students in a timely and on-target base. In this way, students would be able to accumulate their positive attitude towards statistics courses and as a result achieve better learning.

References


