The Impact of Statistical Reasoning Learning Environment: A Rasch Analysis

Shiau Wei Chan*, Zaleha Ismail1, Bambang Sumintono3

1Faculty of Technology Management and Business, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
2Faculty of Education, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
3Institute of Educational Leadership, Universiti Malaya, 59990 Jalan Pantai Baru, Kuala Lumpur, Malaysia

Generally, statistics is taught using the traditional method that only promotes students' procedural knowledge instead of conceptual knowledge. Therefore, a model called Statistical Reasoning Learning Environment (SRLE) was utilized in this study to develop students' statistical reasoning ability. The effectiveness of this model was measured using the Rasch measurement model. In this study, 67 tenth-grade students participated in the quasi-experiment, where 35 students were assigned as the experimental group and 32 were assigned as the control group. The experimental group was treated with the SRLE instruction, while the control group received no treatment. Pre-test and post-test were given to both groups before and after the treatment. Then, the data were analyzed, including the validity and the reliability of the test, person and item reliability, separation, and person-item map. The findings revealed that the statistical reasoning of the experimental group was better than the control group. The SRLE instruction could be applied for future investigations, owing to its great impact on students' statistical reasoning.

Keywords: Statistical Reasoning Learning Environment, Rasch measurement model, Statistical Reasoning

1. INTRODUCTION

To date, statistics has become a mandatory subject from primary schools to universities as it plays important roles in many areas, such as education, engineering, science, economics, business, and so forth. Nevertheless, the instructors are apt to utilize the traditional assessments and instructions, which involve the routine rules, memorization of formula, and computing skills while teaching statistics. Such circumstances could not enhance the students' conceptual knowledge in statistics and have caused students to only comprehend procedural knowledge. Consequently, students have poor statistical reasoning and hold numerous misconceptions related to the statistics, for instance, the measures of the central tendency, variability, and distribution. As far as these matters are concerned, it is proposed that integrating the Statistical Reasoning Learning Environment (SRLE) in the statistics classroom can foster students' statistical reasoning which focuses on the approaches used to reason with the statistical concepts and make sense of statistical information.

The SRLE is a model that has six instructional principles, which are recommended by Cobb and McClain. It is based on the social constructivism theory. It is different from the traditional instruction in terms of the role of technology, focus of lesson, role of teacher, data, discourse, assessment, and function of textbook. Despite the application of SRLE in schools, there have been very few studies that reported on the effects of the SRLE model. Thus, this study was undertaken to bridge this gap. The purpose of this study was to determine the impact of SRLE using the Rasch measurement model.
There were three research objectives that drove this study, i.e.:

a) To identify the validity and reliability of the pre-test and post-test
b) To determine the person and item reliability, and separation in the pre-test and post-test.
c) To examine the differences of statistical reasoning ability between the students who were exposed to the SRLE and the control group in the post-test.

2. STATISTICAL REASONING LEARNING ENVIRONMENT (SRLE)

Statistical reasoning is defined as "the way people reason with statistical ideas and make sense of statistical information. It involves making interpretations based on sets of data, or statistical summaries of data. Students need to be able to combine ideas about data and chance, which leads to making inference and interpreting statistical results (p.101)". As previously stated, there are six instructional principles in the SRLE model, i.e. focusing on enhancing the conceptual understanding of the basic statistical concepts, utilizing real and motivating data, employing classroom activities, incorporating the usage of technological tools, encouraging classroom discourse, and exploiting alternative assessment. In the SRLE class, the central statistical concepts are emphasized, for instance, the measures of the central tendency, variability, and distribution. Generally, there are too many materials to be covered in the Malaysian mathematics syllabus, thus, resulting in superficial understanding of the concepts among the students. In addition, students fail to observe the relationship between various statistical concepts, such as the inability to link the concept of distribution to sampling distribution. Therefore, it has been suggested that the students should be given more emphasis on conceptual knowledge of the basic statistical concepts in order to promote the students' deeper understanding.

Additionally, real and motivating data are encouraged to be used in the SRLE class. Several ways have been proposed by Aliaga et al. to obtain real data, for instance, collecting data which are relevant to students' life in the classroom and searching data from the internet. According to Hall, real data can engage the students actively in the learning and aid them to acquire fundamental statistical knowledge and ability. Not only that, the usage of real data can also raise their desire to learn, as well as diminish their misconceptions.

Besides, the instructors employ carefully designed classroom activities to foster students' statistical reasoning in the SRLE class, for example, through collaboration, inquiry, discussions, and so on. There are many advantages of using active learning, including enhancement of critical thinking, long term retention, and formation of confidence and positive attitudes. Cooperative learning is deemed one type of active learning. It involves students to work together in a group to carry out tasks and achieve expected targets. In cooperative learning, the students are given more responsibility to deal with their own learning. Furthermore, in the SRLE class, technological tools are utilized to enhance statistical reasoning among the students. The rapid evolution of information technology has transformed the education system. During the last few decades, information technology was incorporated into the teaching and learning processes and no longer restricted to the conventional approach in the classroom such as using paper-and-pencil activities. Chan and Ismail asserted that the usage of information technology can enhance students' statistical reasoning. GeoGebra software, a dynamic mathematics software, was used in this study. It can be downloaded freely from the internet. By using GeoGebra, students can produce graphs and visualize the multiple representations of the data.

Moreover, classroom discourse is promoted in the SRLE class. Generally, there is not much discourse in the traditional class as the instructors usually provide the information to the students. However, in the SRLE class students are encouraged to engage in discussions to explicate their reasoning and defend their answers. The students also respond to each other’s questions and learn to ask questions, especially open-ended questions. Such questions allow students to describe their understanding of the learning process, as well as provide understanding in their own words. Hence, the instructors should form a classroom climate where the students feel safe to articulate their opinions.

In the SRLE class, alternative assessment is utilized as the focus of statistics has gradually moved from traditional assessment to alternative assessment. It enables the students to give their reasoning, discuss their way of thinking, and support their judgments and conclusions; examples of alternative assessments are written reports, oral presentations, article critiques, minute papers, projects, reflective journals, and portfolios. Alternative assessments are different from the traditional assessments as they are designed based on the current curriculum, give chance to the students to show their full ability or potential, and assess the learning of the students from different areas.

Social constructivism, which is strongly affected by the work of Vygotsky, is the underlying theory of the SRLE. This theory is defined as a view of learning where the students construct their own knowledge by working together with the instructors and other students. Vygotsky believed that the students cooperate and work together among themselves to achieve a shared goal, and the adult or more talented peers may help them to conduct the tasks within their Zone of Proximal Development. Scaffolding allows instructors to provide students prompt, support, problems, and coaching in order to aid the students before they can complete a task.
independently. The social interaction to construct knowledge depends on one or both of the interaction between teacher and student as well as the interaction between student and student. Such situation can guide the students to a higher level of understanding. In addition, the cognitive development of students is also influenced by the language and culture in the social constructivist learning\textsuperscript{3}. In SRLE classroom, social interaction has taken place and students are allowed to reason their learning and engage in the discussion supported by the technological tool. To determine the effectiveness of the SRLE model, statistical reasoning instruments are created in this study in order to be utilized in the SRLE classroom. In addition, the statistical reasoning ability is investigated between the students who exposed to the SRLE instruction and those who did not exposed to this instruction.

3. METHODOLOGY

3.1 RASCH MEASUREMENT MODEL

In this study, a Rasch analysis was performed using WINSTEPS version 3.73. The Rasch measurement model is a model-based measurement technique, which has gradually become famous for scale creation\textsuperscript{22}. The probability of a student to answer the same item accurately depends on the person's ability and item difficulty\textsuperscript{25}. The Rasch model was employed in this study due to its several advantages\textsuperscript{24}. First, the fit statistics of any item can be identified. Second, the item calibration is not affected by the students' ability and it is sample free. Third, the standard error of calibration can be used to determine the accuracy of every item. Fourth, the item difficulties can be estimated from a variety of samples and changed into a common scale. Fifth, even though the students do not have any item in common, their ability can be also compared as the Rasch model is test-free person measurement. Sixth, Chi-square of person fit can be applied to evaluate the quality of measurement. Finally, the design and construction of test can be managed easily by using the Rasch model\textsuperscript{24}.

3.2 PARTICIPANTS

The research design of this study is a quasi-experimental design, where the individual participants were non-randomly assigned to the experimental group and the control group\textsuperscript{27}. 67 tenth-grade students were involved in the present study. They came from two intact classes in a Malaysian secondary school. The students in these two intact classes were assigned into experimental and control groups. There were 35 students in the experimental group and 32 students in the control group. Ultimately, the number of the participants in the experimental group was reduced to 32 students, as three students were absent from the post-test. The experimental group had 11 males and 24 females. Whereas, the control group comprised of 9 males and 23 females. The age of all the participants was 16 years old. In this study, the experimental group was treated with the SRLE instruction, while the control group was given no treatment. However, both groups were administered with pre-test and post-test before and after the treatment. As regards to ethical consideration, the researcher asked for the students' consent to take part in this study and their data was kept anonymous and confidential.

3.3 INSTRUMENTATION

There were three statistical reasoning instruments used in this study, namely, pre-test, technology-based statistical reasoning tasks, and post-test. The topic covered in these instruments is descriptive statistics, which involves the measures of central tendency, distribution, and variability. There were five tasks with 52 items in the technology-based statistical reasoning tasks. It was utilized in the SRLE instruction to develop the students' statistical reasoning ability. Meanwhile, the pre-test and post-test consisted of 25 true and false problems each; the students had to select true or false in part A and give their explanations in part B. These two instruments were used to evaluate the impact of SRLE on statistical reasoning ability among the respondents. The students were required to select true or false answer options in part A and give their explanation in part B. Each answer was categorized as a dichotomous response, as the items with two potential responses (true or false) were handled by using the Rasch model in the study\textsuperscript{22}.

3.4 PROCEDURE

At the outset of this study, the pre-test was administered to the experimental group and control group. Students had one hour to complete the test. After that, the experimental group was given treatment with the SRLE instruction and the control group received no treatment. In the SRLE class, the researcher focused on the central statistical ideas, i.e. the measures of central tendency, variability, and distribution. Besides, real data were employed in the technology-based statistical reasoning tasks. In addition, cooperative learning was utilized as classroom activities where the students shared laptops in pairs to solve the tasks together. Furthermore, GeoGebra software was the technological tool used in this study. Moreover, the researcher made use of good questions, such as, 'What would happen if...?' and 'What do you think' in order to enhance the classroom discourse. The alternative assessments that were used were the pre-test, post-test, and technology-based statistical reasoning tasks. After the treatment, both groups were given one-hour post-test.

4. FINDINGS

4.1 VALIDITY AND RELIABILITY

In order to measure the validity of the test, fit
statistics was utilized. There were some misfitting items in the pre-test and post-test, as their values of point measure correlation were negative and the outfit z-standard was more than two. Hence, 6 items were removed from the pre-test, i.e. 3a, 3b, 9a, 9b, 17a, and 17b. Meanwhile, 10 items were deleted from the post-test; 6a, 6b, 12a, 12b, 17a, 17b, 20a, 20b, 21a, and 21b.

After those items were eliminated, the reliability of both tests was determined, as displayed in Table 1. The Cronbach’s alpha tells us the internal consistency reliability of the tests. As for pre-test, it was 0.67, which was considered as acceptable. As for post-test, it was 0.88, which was regarded as good. According to Fisher, the person reliability of the pre-test is poor if it is only 0.54. As for this study, the person reliability of the post-test was good, which was 0.81. Moreover, the item reliability for the pre-test and the post-test were 0.93 and 0.92 respectively, which were very good. In addition, the person and item separation for both tests had sufficient spread as they were more than 1.0.

Table 1. Reliability and separation of person and item

<table>
<thead>
<tr>
<th>Test</th>
<th>Cronbach alpha</th>
<th>Person</th>
<th>Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reliability</td>
<td>Separation</td>
<td>Reliability</td>
</tr>
<tr>
<td>Pre-test</td>
<td>0.67</td>
<td>0.54</td>
<td>1.08</td>
</tr>
<tr>
<td>Post-test</td>
<td>0.88</td>
<td>0.81</td>
<td>2.09</td>
</tr>
</tbody>
</table>

4.2 PERSON-ITEM MAP

Figure 1 demonstrates the results obtained by the experimental and control groups in the post-test. The person-item map was utilized to compare the differences between experimental and control groups. The distribution of the students’ ability and item difficulty was positioned on the same logit scale of the person-item map. The ability of the students from both groups is on the left side of the map, whilst the item difficulty is on the right side of the map. Higher logits represent students with higher ability on the left side and more difficult items on the right side and vice versa. It was found that the statistical reasoning ability for the experimental group was more enhanced compared to the control group after the SRLE treatment. This is because more students in the experimental group were located above the average of the test items, while the students in the control group were located below the average of the test items. In other words, the students from the experimental group were capable of solving the test items compared to the control group.

5. DISCUSSION

Based on the findings, the three research objectives of this study have been achieved. As for the first research objective, the validity and reliability of the pre-test and post-test were obtained. Both instruments were judged as acceptable and good. Besides, as for the second research objective, the person and item reliability and separation were attained. The Cronbach’s alpha is unable to identify problem with the person or item, but with Rasch analysis, a problem may be identified by looking at the person and item reliability. As for the third research objective, it was noticed that the statistical reasoning ability of the students who were exposed to the SRLE was better than the students in the control group. Therefore, the SRLE treatment had a great impact on enhancing the students’ statistical reasoning ability. This result was consistent with the study of Sun and Buys, whereby the students’ critical thinking, learning, and research skills were improved after they went through the six principles of STLE treatment, which was similar to the SRLE treatment.

Furthermore, from this study, it was found that SRLE is a powerful instructional model that integrates six principles. By using this model, the students could engage actively in the classroom. However, this model has been undervalued due to a lack in empirical support. In this study, the effectiveness of this model has been confirmed since it improved the students’ learning outcomes. This shall serve as an empirical support to this model and can become a guideline for future studies. Not only that, it is also possible to perform studies, regardless of whether they are related or not related to this subject matter, on different grade levels to further confirm the effectiveness of SRLE. Besides, different types of educational software can be used as the technological tool as well.

On the other hand, the development of the three statistical reasoning instruments in this study provides an alternative teaching method for the instructors. The statistics questions that are usually found in the Malaysian mathematics textbook are close-ended, routine, and traditional. However, it is different with the created statistical reasoning instruments, as they are comprised of open-ended, non-routine, and non-traditional questions. Not only that, the statistical reasoning tasks utilized in the
SRLE instruments involved the usage of technological tools, which is not commonly found in the traditional assessments. Moreover, three statistical reasoning were integrated into the statistical reasoning instruments, including the measures of central tendency, distribution, and variability. Such combination allows the students to understand that these concepts are actually interrelated ideas². In this study, the statistical reasoning instruments had demonstrated their ability to promote and assess the statistical reasoning among the secondary school students. Hence, the instructors and researchers can employ them in the future studies.

6. CONCLUSION

In sum, the findings of this study indicated that the SRLE instruction had positive effects on promoting statistical reasoning among tenth-grade students. Thus, the instructors and researchers should adopt SRLE instruction during in their statistics classrooms. Further explorations of the SRLE instruction can be done in terms of different grade levels, school, culture, gender, background, and so forth.

REFERENCES


