Statistical Graph Interpretation Skills among Pre Service Teachers

Norabiatul Adawiah Abd Wahid, Suzieleez Syrene Abdul Rahim, & Sharifah Norul Akmar Syed Zamri

***Abstract* -- Statistical graphs have been widely used as the medium to spread information to the public. Therefore, the Ministry of Education is emphasizing on this topic as early as Standard 3. The learning process regarding this topic have been continuously done up to university level. There were various parts of study in statistics lesson with one of it was on statistical graph which mostly focused on the interpretation skills. Therefore, statistical graph interpretation skills among the teachers have been highly demanded to help them while conducting their lessons on this topic. This study had been conducted to look at how pre service teachers interpret a bar graph represented to them. At the same time, we have looked at how the participants did their interpretations based on the mean value which was also given in the questions. In order to fulfil our objectives we conducted 30 to 45 minutes interview sessions with five participants. All of our participants are pre service teachers who are at their last year of their studies at one of teachers training college (IPG) in Malaysia. An interview session was done personally with each one of them. Based on our findings, we have determined that there are four graphs interpretation skills that had been used when our participants did their interpretation of a bar graph given. We have also determined that the participants of this study have understood the effect of mean value to the set of data represented by the given bar graph. This proves that all of them did not face too much problem while doing the graph interpretations. At the same time, our participants could also determine the impact of outliers to the mean value of the data.**

***Keywords* -- Statistical Graph, Graph Interpretations Skills, Mean Value, Outliers**

Norabiatul Adawiah Abd Wahid*, Suzieleez Syrene Abdul Rahim, & Sharifah Norul Akmar Syed Zamri, Faculty of Education, University of Malaya, Kuala Lumpur (e-mail*: norawahid@siswa.um.edu.my).

I. INTRODUCTION

Nowadays, the needs for teachers to comprehend statistics and acquire all the skills related to statistics is a must. In Malaysia, statistics has been listed as one of four learning areas in mathematics curriculum for primary school pupils. The importance of this learning area have been proved when lessons on statistics have been implemented in national curriculum beginning from Standard 3. Lessons on statistics starts by focusing on data handling. Meanwhile, lessons on statistical graphs have been introduced in Standard 4 when pictographs, bar charts and pie charts have been the focus of study in Mathematics Standard 4 syllabus. Lessons on statistics was continuously implemented in Malaysia curriculum until university level (Chan & Ismail, 2012). Therefore, the success for this learning area will mostly depend on school teachers competencies in statistics skills.

The implementation of learning area regarding statistics as early as Standard 3 and statistical graphs starting from Standard 4 has proved how important statistics is. It is increasingly important especially when most of the information nowadays are represented by various types of statistical graphs. By using various types of statistical graphs, it helps to facilitate the delivery of information to the readers. This has brought on the needs of statistical graph interpretation skills to help on interpreting various types of statistical graphs. But, the problem arises when interpretation of statistical graphs might be considered as an activity which is related to a complex range elements and processes (Carlos Monteiro & Ainley, 2007).

Therefore, teachers should ensure that they have comprehended all the skills and processes first before they begin their lesson on this learning area. But based on the research findings, researchers have found that there were much problems arising among the teachers. Therefore, in this study we will look at how the pre service teachers who took part in this study interprets a bar graph which has been presented to them. We will look at the skills used to interpret a bar graph given. At the same time, we will also look at how they interpret the meaning of mean value that have been given in the question and the impact of it on their interpretations on the bar graph.

II. RESEARCH PURPOSES

This research would try to identify the skills used by the participants of this study when they interpret a bar graph. Besides that, this study would look at how the participants interpret the meaning of mean value based on the data given in the bar graph.

III. LITERATURE REVIEW

Researches which have been done based on statistical learning are still in early stages (Leavy, 2010). Therefore we have done our research by focusing on this topic of study especially among pre service teachers. The focus of our study was on the skills used by our participants when they carried out their interpretations on the bar graph given.

There was a prove that the skills to interpret sets of data are not only limited on how the interpreter used their statistical knowledge. According to Pierce, Chick, Watson, Les and Dalton (2014), it also requires an ability to think critically with data. This ability have been the main interest for researcher to do their study in the past few years. It is known as statistical literacy.

The idea on statistical literacy have been introduced by Gal (2002). Starting from the introduction of statistical literacy by Gal (2002) there were many researchers who put their interest on this field of study afterwards. Therefore, the issues studied on statistical literacy have

been expanded by various researchers. The subsequent researchers have tried to look at how the data interpreter interprets their data. At the same time, interest on how reading and reasoning process on statistical graphs have also been growing (Wang et al., 2012). Research conducted by Shah and Freedman (2011) found that there were three factors affecting how a person interpret bar and line graphs. Those three factors were how the readers of both graphs done their interpretations as a function of the format, the reader familiarity with the content depicted in a graph and the readers graphicacy skills.

Skills on graph interpretations have been introduced by Friel, Curcio and Bright (2001). It is known as graph sense. There were six behaviours or skills should be shown by the graph readers when they interpret a graph. Those six skills were:

a) To recognize the components of graphs, interrelationship between the components and effect brought by all of the components on the presentation of information in graphs.
b) To speak the language of specific graphs when reasoning about information are displayed in graphical forms.
c) To understand the relationships between a table, a graph, and the data that had been analysed.
d) To respond to different level of questions that relates to graph comprehension.
e) To recognize when one graph are more useful than another on the basis of judgement task involved and the kinds of data being presented.
f) To be aware of one's relationship to the context of graph, with the goal to interpret and making sense to what been presented by a graph and avoid personalization of the data.

Shoughnessy (2007) have added two more skills to the above list. The skills that have been added by Shoughnessy (2007) were:

a) To determine the relationship between variables in a set of data.
b) To determine the factors affected variations in a set of data.

Even though various suggestions have been listed, problems still arise. Monteiro and Ainley (2007) have found that some of the primary school teachers who took part in the study did not know basic notions related to graphs. This issue is parallel with the findings that have been done by Glazer (2011). Glazer (2011) found that pre-service teachers are not prepared to teach topics involving data analysis and graph comprehension. At the same time, he realized that some pre-service teachers who were involved in his study considered graph as a picture or map when they do their interpretation. These findings brought to another difficulty when the pre-service teachers tried to do their interpretation on statistical graphs. This is proved by the findings by Koleza and Kontogianni (2013). Koleza and Kontogianni (2013) found that pre-service teachers who took part in their study could not give proper justifications to statistical claims in the context of social discussion. At the same time, Koleza and Kontogianni (2013) found that their respondents faced difficulties to master three types of knowledge needed to help them to understand statistical graphs. The knowledges were context knowledge, literacy skills and critical stand.

Issues arising from the various findings regarding graph interpretation have been found in studies on the other statistics skills too. A study by Ismail and Chan (2015) found that the participants of their study were not able to give explanation on how they solved the questions involving measures of central tendency. The main reason behind this problems was the tendency by the participants to solve problems procedurally and by memorizing the formula without understanding the concepts. These findings have also been found by Monteiro and Ainley (2007), Glazer (2011) and Koleza and Kontogianni (2013) in their study which were discussed earlier.

Therefore, by conducting this research, we hope to determine either the same problems as mentioned by the other researchers above could also be found in our study. If the problems arise, we hope that those findings could be used as the guidelines for the lecturers and IPG Malaysia to help them to strengthen statistics courses offered to pre-service teachers as per mentioned by Dogan Temur, Akbaba Dag, and Turgut (2015).

IV. METHODOLOGY

This study is a case study. Five pre-service teachers have been involved as the participants for this study. To easily discuss the findings, all of them have been coded. They have been coded as P1, P2, P3, P4 and P5. All of our participants are final year students at one of the teachers training college (IPG) in Malaysia. After they graduated from their study, they will be posted to primary schools to teach Mathematics. Throughout their study at IPG, they have been in two courses directly related to statistics. Those two courses are Statistik (MTE3053) and Mengajar Geometri, Ukuran dan Pengendalian Data (MTE3113). When this study was conducted, all of our participants have completed both courses.

For the purpose of this study, a bar graph, adapted from PISA 2003 Web Sample (Fig. 1), has been use as the instrument. Besides that, mean value for both groups which data have been represented by the bar graph have also been given as 62.0 for Group A while 64.5 for Group B.

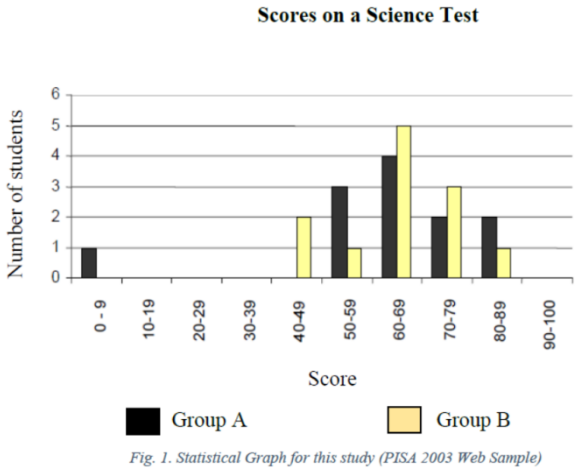

Fig. 1. Statistical Graph for this study (PISA 2003 Web Sample)

All of the participants in this study have been asked to interpret the bar graphs given as the above Figure 1. All of the participants have been allowed to have some time to interpret the bar graphs given before the interview started. An interview was carried out for 30 to 45 minutes with each participant individually. Questions asked throughout the interview session were based on the response given by the participants. The purpose of the interview was to get views on their opinions and interpretations of the bar graph. Bar graph in Fig. 1 were given to all the participants before interview session began. By doing this, the participants would have some time to interpret the graphs and get some views about the information represented by the bar graph before the discussion begins.

V. FINDINGS

Based on our findings, we determined that our participants had used five different types of graph interpretations skills when they interpreted the given bar graph. The five skills determined were based on the suggestions by Friel, Curcio and Bright (2001) and Shoughnessy (2007). We have altered the types of skills listed by those researchers to best suit our findings. At the same time we have also looked at how our participants determine the effect of mean value to the set of data represents by the bar graph.

Determine the Factors Affected Variations in a set of Data

Based on our findings, there were three factors affecting the variations in a set of data represents by the bar graph. Those three factors were value of mean, types of exercise received by the students and the backgrounds of each students.

Mean Value

P1, P2 and P4 stated that mean value given in the question have left an effect on the variations of data represented by the bar graph. P1 stated that "If the mean value of certain group is high than that group perform better if mean value is low the achievement of that group is lesser."P1 agreed that a group with high mean value had performed better compared to the group with lower mean value. This opinions were agreed by P2. In the same time, P2 also mentioned the effect of the outliers as a reason why the mean value of Group A is lower than Group B. He mentioned this issue as "The mean value of (Group B) is higher...If we look at this group (Group A) there was a student with the lowest mark here." Meanwhile P4 mentioned that the achievements by the two groups are not much different between each other. Quite similar with P2, P4 also stresses on the effect brought by the number of failures in each group as affecting the mean value. P4 stated "But both groups of students have high achievements. This is because there were not much difference in their mean value. It is only 2.5 difference between each group. The different was on the number of failures in each group. It affected the mean value."

Types of Exercise Received

Three of the participants agreed that variations in sets of data by those two groups happened because they received different types of practise questions. At the same time, they have agreed that practise questions provided after each failure could help them to pass the examination. P1 mentioned "These two values whom failed the examinations could pass if given more practice questions. This is because the difference is only 10 marks. But, it will be difficult for one of the student from Group A to pass the examination as his marks was at the range of 0 to 9." Meanwhile, P4 stressed on how the failure could help themselves to pass the examination. He stated "students who gets around 48 or 49 marks could pass the examination if they work harder. Based on the teacher's opinion, Group B is better than Group A because the marks of the students from Group B are closer to the passing marks. If they work harder, they could actually pass the examination. A student from Group A needs to work harder." Besides that, P5 have mentioned on how teachers can play their role on helping their students to get better results. "These two students can get better results if their teacher helps them. Yet, it will be more difficult with Group A compared to Group B. Comparatively, the difference of marks of Group A to the passing marks is closer compared to the marks of Group A. The teacher will find it easier to help students from Group B as the ones who failed are not really weak as compared to the ones in Group A. Teacher should be more focused on the failed students of Group A."

Background of Each Student

P5 is the only participants who mentioned about the students' background as the factor why there were variations in the sets of data. "The failed students might not have failed due to being weak in his studies. There could be other factors involved such as rarely attending classes, coming from poor family with financial issues or involved with any accident."

Determine the Relationship between Variables in a set of Data

The bar graphs given represents achievement for two different groups in a Science test. There were two variables based on the bar graphs. The variables were student marks and the number of students. Our participants have been able to relate both variables when they discussed the information represented by the bar graph. All the participants have put their interpretations on achievement to best describe the bar graph. For P4 she have been able to describe that the overall achievement for all of the students as "average students". P4 mention "Most of the students from both group achieved 60 to 69 marks. Most students were in this range of mark which make them average students." Meanwhile P5 divided his answer into two different categories when the researcher ask on which group has the better performance. The two categories were the mean value and the number of students in different range of marks. "Group A is better if we look at the number

of students who passed the test. If we look at the value of mean given for Group A 62 B 64. Based on highest passing marks, Group A could be chosen, but I will choose Group B if it is based on mean."

Effect Brought by all of the Components on the Presentation of Information in Graphs

By looking at the bar graph, there were outliers which could play a large impact to the set of data represented. The outlier data from Group A has been determined as the effect why Group B's achievements was greater than Group A. This have been stressed by P1 who stated "Only one student from Group A affected the results of that group because he only gained around 0 to 9 marks. That affected the overall results for Group A." The outliers also have played some role on why the mean value of Group A is lesser than Group B. P2 mentioned about this effect in his answer. "Because their score is the lowest so it affected their mean value." P5 mentioned "Group A's mean value is lower because they have a student with marks less than 10 which has affected the overall group marks."

Suggestion for Other Graph

One of the interpretation skills that have been listed by Friel, Curcio and Bright (2001) was the ability to recognize when one graph are more useful than another on the basis of judgement task involved and the kinds of data being presented. P3 gave his suggestion to use separate bar graph to represent the achievements of each groups. As mentioned by P3 "have separate graphs for Group A and Group B. This would be better as the students can look at their own group's graph. They are still allowed to see the other Group's graph too."

VI. DISCUSSION AND CONCLUSION

Based on our findings, we have determined four different graphs interpretation skills that have been used by our participants while interpreting the given bar graph. The first skill was the ability to determine the factors that made variations in the set of data. P1, P2 and P4 mention about the effect of mean value to the variations in the set of data represented by bar graphs. Besides that, our participants have been able to determine the effect of outlier in Group A as the reason why their mean value is slightly lower than Group B. This findings are contradicting with the findings of the study done by Chick and Pierce (2008). Chick and Pierce (2008) found that the participants in their study do not make any attempt to discuss on what calculations such as the mean actually reveal and how they relate to the data.

Besides that, the bar graph used in this study is quite related with participants' background as pre-service teachers. Bar graph provided in this research discussed on the achievements of two groups of students in Science test. Based on participants' experience, they have been at schools for their internship twice. Based on our discussion with the participants, some of them have experienced marking the exam paper. Therefore, what was represented by the bar graph were quite related with their experience and also relates with their future task. Therefore, there were needs for participants to discuss on the information represented by the bar graphs without much problems to deal with. However, our participants faced difficulties to discuss in detail about the information given by the bar graph. The discussion that they made during the interview session was not in depth enough. These findings were in line with the findings by Koleza and Kontogianni (2013). Koleza and Kontogianni (2013) found that the pre-service teachers who took part in their study were able to read and interpret simple graphs given to them. But, they faced difficulties when they were needed to give proper justification in the context of social discussion.

Even though our participants were not able to discuss in detail about their interpretations, all of them have been able to discuss the relationship between all of the variables. They have also described the achievements for both group of students as average. This has brought to the ability on determining the effect brought by all of the components on the presentation of information in the bar graphs. The most discussed point on this matter was on the outlier happened in Group A. P3 and P4 also gave some suggestions on the other types of graph that could represent the same set of data as the bar graph.

Even though the participants can do their interpretations on the given bar graph, research on other types of statistical graphs should be carried out. This is to make sure that the participants of upcoming study will be able to use all of the graph interpretation skills while interpreting different type of graphs. Besides that, a study on the effect of other types of central tendency (median and mode) to a set of data should also be conducted.

VII. REFERENCES

Chan, S. W., & Ismail, Z. (2012). The Role of Information Technology in Developing Students' Statistical Reasoning. *Procedia - Social and Behavioral Sciences*, *46*, 3660–3664. http://doi.org/10.1016/j.sbspro.2012.06.123

Chick, H. L., & Pierce, R. (2012). Teaching for statistical literacy: Utilising affordances in real-world data. *International Journal of Science and Mathematics Education*, *10*(2), 339–362. http://doi.org/10.1007/s10763-011-9303-2

Chick, H. L., & Pierce, R. U. (2008). Teaching Statistics at the Primary School Level: Beliefs, Affordances, and Pedagogical Content Knowledge. *Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. Challenges for Teaching and Teacher Education*, 1–6.

Dogan Temur, Ö., Akbaba Dag, S., & Turgut, S. (2015). Some Reflections From Pre-Service Elementary Teachers' Practice Teaching on the Area of Understanding Data in the Math-Teaching Course. *International Electronic Journal of Elementary Education*, *7*(3), 355–370. Retrieved from https://acces.bibl.ulaval.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=eue&AN=108357550&lang=fr&site=ehost-live

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: critical factores influencing comprehension and instituctional implications. *Journal for Reearch*

in Matemátics Education, *32*(2), 124–158. Retrieved from http://filebox.vt.edu/users/sboyce/Curriculum Summer 03/Stat Articles/JRME2001-03-124a.pdf

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, *70*(1), 1–51. http://doi.org/10.1111/j.1751-5823.2002.tb00336.x

Glazer, N. (2011). Challenges with graph interpretation: a review of the literature. *Studies in Science Education*, *47*(December), 183–210. http://doi.org/10.1080/03057267.2011.605307

Ismail, Z., & Chan, S. W. (2015). Malaysian students' misconceptions about measures of central tendency: An error analysis, *93*, 93–100. http://doi.org/10.1063/1.4907430

Koleza, E., & Kontogianni, A. (2013). Assesing statistical literacy: What do freshmen know? In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education* (pp. 840–849). Antalya, Turkey: Middle East Technical University and ERME.

Leavy, A. (2010). The challenge of preparing preservice teachers to teach informal inferential reasoning. *Statistics Education Research Journal*, *9*, 46–67. Retrieved from http://www.stat.auckland.ac.nz/~iase/serj/SERJ9(1)_Leavy.pdf

Monteiro, C., & Ainley, J. (2007). Investigating the interpretation of media graphs among student teachers. *International Electronic Journal of Mathematics Education*, *2*(3), 188–207.

Monteiro, C., & Ainley, J. (2007). Investigating the Interpretation of Media Graphs Among Student Teachers. *International Electronic Journal of Mathematics Education*, *2*(3), 187–207.

Pierce, R., Chick, H., Watson, J., Les, M., & Dalton, M. (2014). A statistical literacy hierarchy for interpreting educational system data. *Australian Journal of Education*, *58*(2), 195–217. http://doi.org/10.1177/0004944114530067

Shah, P., & Freedman, E. G. (2011). Bar and Line Graph Comprehension: An Interaction of Top-Down and Bottom-Up Processes. *Topics in Cognitive Science*, *3*(3), 560–578. http://doi.org/10.1111/j.1756-8765.2009.01066.x

Wang, Z. H., Wei, S., Ding, W., Chen, X., Wang, X., & Hu, K. (2012). Students' Cognitive Reasoning of Graphs: Characteristics and progression. *International Journal of Science Education*, *34*(January 2015), 2015–2041. http://doi.org/10.1080/09500693.2012.709333