Singapore Secondary School Students' Understanding of Statistical Graphs

Wu, Yingkang
National Institute of Education, Nanyang Technological University, Singapore
e-mail: ykwu@mail.nie.edu.sg

Abstract: This study investigated Singapore secondary school students' understanding of statistical graphs. Students' Understanding of Statistical Graphs (USG) was defined from four aspects: graph reading, graph interpretation, graph construction and graph evaluation. Altogether 907 students (13 to 15 years old) completed a Test of students' Understanding of Statistical Graphs (TUSG) and 63 of the students were interviewed. The result of a three-way MANOVA shows how students' USG in the four aspects differed with stream, grade level, and gender. A Rasch analysis provides information about the difficulty levels of the items in TUSG. Twelve error categories were identified. The conclusions of the study and implications for teaching and curriculum development are briefly discussed.

Introduction

Compared to the intensive studies on probability, data handling is just an emerging research topic. The development of graphical understanding in students has been identified as one of the three research themes on data handling by Shaughnessy, Garfield and Greer (1996). A number of studies are beginning to shed light on students' understanding of graphs in general and of particular types of graphs. However, most of the studies seem to concentrate on students' ability to read and interpret graphs (Curcio, 1987; Friel & Bright, 1996; Mellor, 1991; Watson & Moritz, 2001). Only a few empirical studies examine graph construction tasks (Graham, 1989; Padilla, McKenzie & Shaw, 1986). With the frequent usage of graphs in mass media, the ability to evaluate a statistical graph becomes increasingly important. Even fewer studies address this issue (Mooney, 1999; Watson, 1997).

This study investigated students' ability to read, interpret, construct and evaluate statistical graphs concurrently. It intended to depict a comprehensive picture of secondary school students' understanding of statistical graphs in Singapore, where statistical graphs form the main focus of the statistics topic in the lower secondary mathematics curriculum (Ministry of Education, Singapore, 2001). More specifically, this study sought to answer the two research questions:

1. What is Singapore secondary school students' understanding of statistical graphs?
2. What are the errors students make when working with different types of statistical graphs?

A conceptual framework of Understanding of Statistical Graphs (USG)

In this study, students' Understanding of Statistical Graphs (USG) was defined from four aspects: graph reading, graph interpretation, graph construction and graph evaluation. Each aspect has several components. Each component under graph reading, graph interpretation and graph evaluation has two types: type (a) and type (r). Type (a) asks for final answers, such as values and choices. Type (r) requires descriptions, explanations, justifications, and reasoning to support the final answers. The definitions of each aspect and its components are shown in Figure 1. This conceptual framework characterizes the nature of USG.

Type (r) is included in the framework for two reasons. First, to use mathematical language to communicate mathematical ideas and arguments precisely, concisely and logically is one of the aims of school mathematics education (Ministry of Education, Singapore, 2001). It is of interest to see how well students could express their mathematical thinking about statistical graphs. Second, type (r) can provide useful information about how students solve the graphical problems and the errors they make.

Methodology

Altogether 907 Secondary 1, 2 and 3 students (13 to 15 years old) in the four streams[1] from five schools participated in the study during February to April in 2003. They completed a Test of

[1] At the secondary level, Singapore students are streamed into one of the four courses (Special, Express, Normal Academic and Normal Technical) based on their PSLE (Primary School Leaving Examination) results.

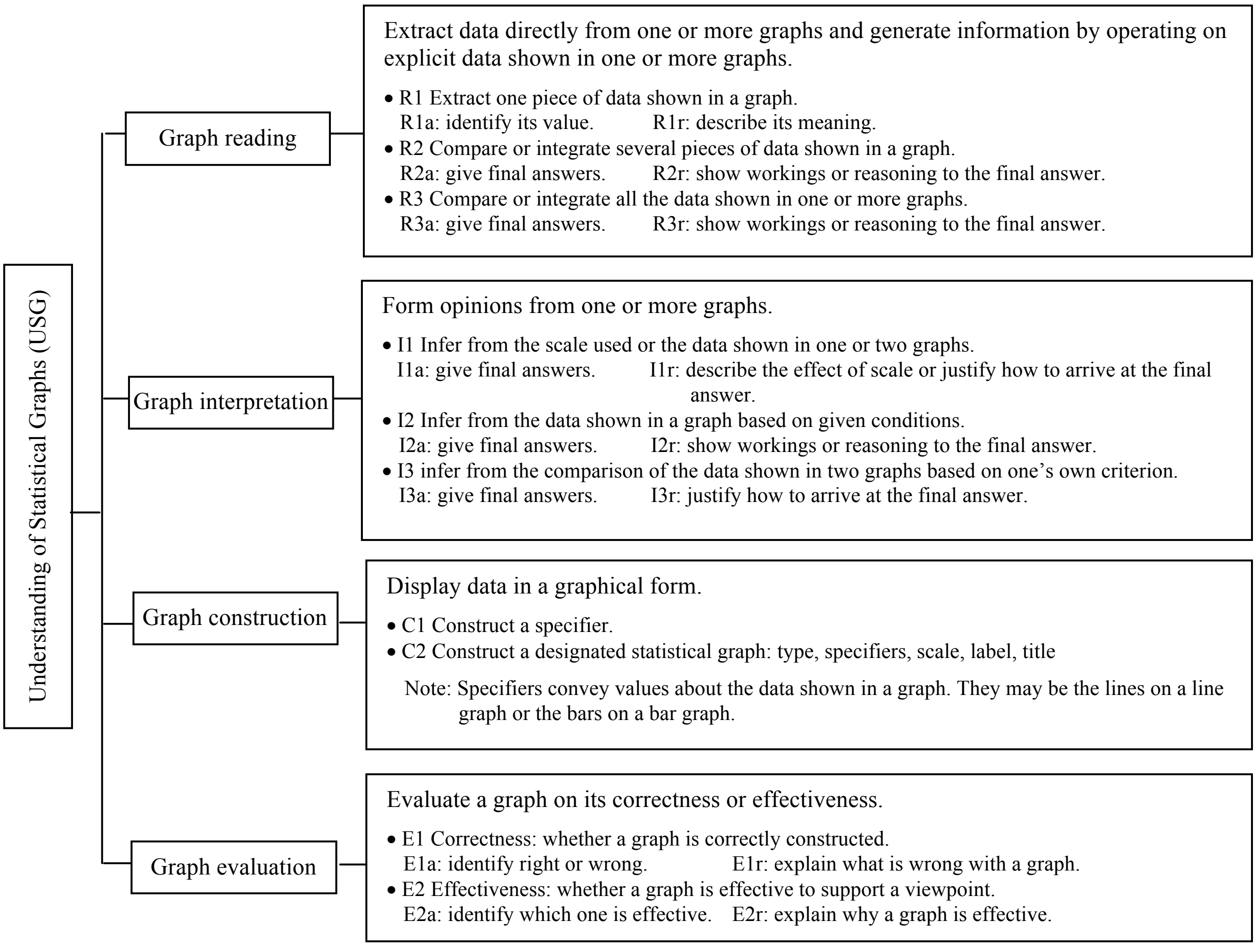

Figure 1. The framework of USG

students' Understanding of Statistical Graphs (TUSG) in about fifty minutes. TUSG consists of 10 questions with 53 items[2], covering six different types of statistical graphs: bar graph, pie chart, line graph, pictogram, dot diagram and histogram. These statistical graphs were chosen because they are within the Singapore lower secondary mathematics syllabus common to all the four streams (Ministry of Education, Singapore, 2001). Sixty three students were interviewed to make clear their thinking processes and the errors they made in solving the problem.

A correct response to each item was scored 1 mark, giving a total of 53 marks for TUSG. The Cronbach's alpha for the whole test was 0.92, showing satisfactory internal consistency. The types of errors students made were coded. An inter-coder reliability of 94.2% agreement between two different coders was obtained.

Results

1. What is Singapore secondary school students' understanding of statistical graphs?

The summary of students' performance in TUSG is shown in Table 1. On average these students were able to solve more than 50% of the items in TUSG correctly but with a wide variation. The four aspects ranked in a descending order of the percent mean scores are: graph construction, graph reading, graph evaluation and graph interpretation. It shows that the students had more difficulties with graph interpretation and evaluation tasks than graph construction and reading tasks. This is expected because the graph interpretation and evaluation tasks are more challenging. First, these tasks go beyond a graph, requiring students to make inferences from the information shown in a graph or apply their graphical knowledge to a new situation, while the graph reading

[2] Each question in TUSG has several parts. If a part of a particular question requires a final answer and the reasoning to the answer, it was treated as two separate items: type (a) and type (r), according to the framework of USG.

and construction tasks are generally limited to a graph itself, asking students to read off the data from a graph, make simple operations on the explicit information shown in a graph, or display the given data in a graphical form. Second, several graph interpretation and evaluation items require students to give explanations or reasoning to justify their answers. Many students were not familiar with these justification questions and sometimes they had language difficulties to express themselves clearly. In contrast, no items under graph construction and relatively fewer items under graph reading require explanations.

Table 1
Summary of Students' Performance in TUSG

Aspect	Total possible score	Raw score Mean	Raw score SD	Percent score Mean	Percent score SD
Overall	53	28.0	10.3	52.8	19.5
Graph reading	22	12.8	5.3	58.1	24.0
Graph interpretation	14	5.9	3.3	41.8	23.3
Graph construction	7	4.4	1.8	63.2	26.1
Graph evaluation	10	5.0	1.6	49.5	15.9

A Rasch analysis was applied to students' raw scores using the RUMM 2020 program. The person separation index was 0.92 and the power of test-of-fit was identified as excellent. The students' abilities and the item difficulties were placed on the same metric, as shown in Figure 2.

```
-----------------------------------------------------------------------------------------------
LOCATION          PERSONS     ITEMS
-----------------------------------------------------------------------------------------------
                         |
  4.0                    | q5r(E1r)
                         | q6e(E1r)
                         |
                       X |
                         | q7a(R1r)
  3.0                    |
                      XX |
                     XXX |
                         | q6b(E1r)
                    XXXX | q8cr(I1r)    q10dr(I2r)
  2.0          XXXXXXXXX |
                   XXXXX | q3b(I1r)
                    XXXX | q9c(I1r)
                  XXXXXX | q8dlabel(C2)
         XXXXXXXXXXXXXXX | q10ca(R3a)
  1.0         XXXXXXXXXX | q9b(I1r)
         XXXXXXXXXXXXXXX | q6d(E1r)     q10cr(R3r)
         XXXXXXXXXXXXXXX | q8ca(I1a)
            XXXXXXXXXXXX | q6c(E1r)     q8dtitle(C2)  q1cr(I2r)   q10br(R2r)   q10ba(R2a)
                 XXXXXXX | q7cr(R3r)    q4cr(I1r)     q7dr(R2r)   q7br(R2r)    q10e(I3r)
  0.0      XXXXXXXXXXXXX | q1ca(I2a)    q10da(I2a)
         XXXXXXXXXXXXXXX | q7ba(R2a)    q4br(R3r)
           XXXXXXXXXXXXX | q9ar(E2r)    q2ba(R2a)
         XXXXXXXXXXXXXXX | q2br(R2r)    q4ca(I1a)     q8br(I1r)   q4a(C1)
           XXXXXXXXXXXXX | q7ca(R3a)    q7da(R2a)     q4ba(R3a)
 -1.0          XXXXXXXXX | q8dtype(C2)  q10a(R1a)
                 XXXXXXX | q3aa(R2a)    q8dscale(C2)  q2ar(R2r)
             XXXXXXXXXXX | q5a(E1a)     q8dbars(C2)
                   XXXXX | q2aa(R2a)    q3ar(R2r)     q1d(E1a)
              XXXXXXXXXX | q8ba(I1a)
 -2.0               XXXX |
                   XXXXX |
                       X |
                       X | q9aa(E2a)    q1a(R1a)
                       X | q6a(E1r)
 -3.0                    | q1b(C1)      q8a(R1a)
                         |
-----------------------------------------------------------------------------------------------
```

Note: 1. X = 4 Persons
2. Each item is noted in the form of qxyz (z is optional). qxy refers to part y of question x, and z could be a or r (a refers to answer and r refers to reasoning to the answer). e.g. q1ca refers to the item asking for the answer to part c of question 1. In particular, for part d of question 8 (it asks students to draw a bar graph) z is one of the five examined points: type, bars (specifiers), scale, label and title.
3. The type of an item is in the parentheses.

Figure 2. Scale of students' abilities and item difficulties

The hierarchy of the 53 items shown in Figure 2 reveals that the difficulty level of items related to two aspects. First, it related to the type of graphs used. Question 10 is about a histogram. Most items under question 10 had positive locations, indicating that they were relatively difficult. Question 1 is about a bar graph and question 2 is about a pie chart. Most items under question 1 and 2 had negative locations, indicating that they were relatively easy. The bar graph and the pie chart represented discrete values, while the histogram described the frequency of the data value rather than the value itself. This could be used to explain why the histogram was more difficult than the bar graph and the pie chart. Second, the item difficulty level related to the type of questions asked. Graph reading and construction items were found generally easier than graph interpretation and evaluation items as suggested by their locations. This is consistent with the pattern shown by the students' percent mean scores in the four aspects (see Table 1). In addition, type (a) items (requiring final answers only) were generally easier than type (r) items (requiring descriptions, explanations, justifications and reasoning). During the interviews, some students mentioned that they were seldom asked to explain or justify how they arrived at their answers in solving mathematical problems. That these students did not have enough exposure to mathematical communication tasks to some extent increased the difficulty of type (r) item.

A three-way Multivariate Analysis of Variance (MANOVA) was conducted to examine how students' USG in the four aspects differed with student characteristics (stream, grade level and gender) using the GLM analysis of SPSS for Windows Release 11.0.1. A multivariate approach was chosen in order to control the Type I error and to detect combined differences that may not be evident from univariate tests (Hair, Anderson, Tatham & Black, 1995). The Bartlett's test of sphericity had a significance level at 0.001 (χ^2 =1456.7, df =9, p<0.001), satisfying the necessary requirement of intercorrelation to justify MANOVA. The Pillai's criterion was chosen for significance testing because it is more robust (Hair et al, 1995, p. 278). Table 2 shows the result of the multivariate test.

Table 2
Results of the Three-way MANOVA Test

Source	Pillai's Trace	F value	p value
Stream	.684	65.08	.000
Grade level	.151	17.97	.000
Gender	.022	4.99	.001
Stream x Grade level	.056	2.07	.002
Stream x Gender	.010	0.72	.735
Grade level x Gender	.007	0.78	.624
Stream x Gender x Grade level	.026	0.97	.499

The three main effects: stream, grade level, gender, and the interaction effect stream x grade level were statistically significant. In general upper grade level students did better than lower grade level students and the performance of the four streams ranked in a descending order was Special, Express, Normal Academic and Normal Technical. The significant interaction effect indicates that the effect of grade level was modified by the effect of the stream. Thus, the following discussions focus on the main effect of gender and the interaction effect of stream x grade level.

In terms of the main effect of gender, the tests of between-subjects effects show that the male students did statistically significantly better in graph reading than the female students did (F(1, 883)=6.41, p=0.012), while the female students performed statistically significantly better in graph construction than the male students did (F(1, 883)=4.82, p=0.028). This result was inconsistent with some previous studies (Curcio, 1987; Lenton, Stevens, & Illes, 2000). No statistically significant differences were found in graph interpretation and evaluation aspects between the male and female students.

With respect to the significant interaction effect between stream and grade level, the tests of between-subjects effects show that it was the graph interpretation (F(6, 883)=4.97, p<0.001), graph construction (F(6, 883)=3.51, p=0.002) and graph evaluation (F(6, 883)=2.44, p=0.024) that were responsible for the overall significant interaction. It indicates that the effect of grade level on students' performance in these three aspects was not the same for all the four streams. Posterior comparisons revealed that the effect of grade level in these three aspects was statistically significant for Express and Normal Academic stream students but not for Special and Normal Technical stream students.

There was no significant interaction effect in graph reading aspect (F(6, 883)=1.92, p=0.076). The main effects of stream and grade level produced the statistical significance in this aspect. The Scheffe post hoc tests show that the mean scores in this aspect differed statistically significantly at the 0.01 level between any two of the four streams and any two of the three grade levels.

2. What are the errors students make when working with different types of statistical graphs?
Many errors were observed in the study and they were summarized into 12 categories. Table 3 shows the error categories and the percentages of the students who made each error category at least once from each grade level and the whole level.

Table 3
Percentages of Students with Each Error Category

Error category	Sec 1 (n=242)	Sec 2 (n=345)	Sec 3 (n=320)	Overall (n=907)
E1 Comprehension errors	69.8%	69.3%	58.1%	65.5%
E2 Unclear explanation errors	42.6%	46.1%	48.4%	46.0%
E3 Computation errors	14.9%	22.3%	19.7%	19.4%
E4 Scale errors	88.4%	93.9%	86.9%	90.0%
E5 Errors related to title, label, specifiers and graph type	98.8%	98.8%	97.5%	98.3%
E6 Pie chart errors	12.4%	18.3%	14.7%	15.4%
E7 Size pictogram errors	69.8%	73.9%	72.2%	72.2%
E8 Appearance-similar but nature-different graphs errors	54.5%	49.3%	32.8%	44.9%
E9 Confusion between frequency and data value errors	2.9%	2.3%	2.5%	2.5%
E10 Errors related to the use of the given information	85.1%	91.3%	85.9%	87.8%
E11 Errors related to the use of context	71.5%	75.4%	60.6%	69.1%
E12 Miscellaneous errors	79.3%	78.8%	73.8%	77.2%

Category E1 is about students' comprehension skills. Around 65.5% of all the students did not understand what a particular question asked for at least once. Secondary 3 students made relatively fewer E1 errors than Secondary 1 and 2 students. Category E2 associates with students' communication skills. Around 46.0% of all the students were observed not able to explain their thinking clearly. The percentages across the three grade levels should be interpreted cautiously. On average more Secondary 3 students (82.8%) and Secondary 2 students (80.9%) attempted type (r) items (requiring description, explanations, justifications and reasoning) than Secondary 1 students (72.8%). Thus, Secondary 2 and 3 students were more likely to make E2 errors than Secondary 1 students. Category E3 refers to a calculation error occurred in performing a correct operation. Overall less than 20% of the students made this type of errors. Comprehension, communication and computation skills are the pre-requisites to solve a graph related problem. The figures presented indicate that students' failure to solve such a problem, to a certain degree, was due to their inability in these skills.

Categories E4 to E7 involve conventions of graphs. About 90% of the students did not know one or more of the following scale conventions: the scale should start from zero in a bar graph, the scale interval should be equal, and the scale should be marked in drawing a statistical graph. Almost all of the students (98.3%) did not know that a graph needs a title and/or the axes in a graph need to be labeled. About 72.2% of the students had difficulty with a size pictogram, which uses different sizes of a picture to represent different quantities. They did not know that sizes of a picture could represent quantities, a pictogram is different from an artistic picture, or a three-dimensional pictogram could exaggerate the intended information dramatically. Compared to the larger percentages of E4, E5 and E7, relatively fewer students (around 15%) made E6 errors. These students did not know how to convert degree to percentage in a pie chart and/or the total of a pie chart is 100%.

Category E8 means that students mistook a graph as another appearance-similar but nature-different graph. For example, many students took the histogram in question 10 as a bar graph in solving the relevant problems. Nearly 45% of all the students made E8 errors. The percentages of the students with E8 errors decrease obviously with the increase of the grade levels, implying that students' ability to differentiate graphs with similar appearances increase with age.

Category E9 is about frequency and data value. Almost all the students were able to distinguish between data value and its frequencies. Although some previous studies reported that students had difficulties in this regard (Friel & Bright, 1996; Hawkins, Jolliffe, & Glickman, 1992), the lower error percentages suggest that it was trivial for these Singaporean students.

Category E10 deals with students' use of the given information to solve a graph related problem. About 87.8% of the students made E10 errors. They were found to select inadequate information from a graph, misused the graphical features (e.g. used the actual heights of the bars when the differences between the heights of the adjacent bars were needed to answer a bar graph related question), and applied incorrect algorithm or invalid reasoning to solve problems.

Category E11 is about students' misuse of contextual knowledge to solve a graph related problem. Nearly 70% of the students were observed to rely on their real life knowledge or their personal experiences at least once to answer a particular question, which is considered as irrelevant from a statistical point of view (Watson & Pereira-Mendoza, 1996). The younger students seemed to have a greater tendency to make E11 errors than the older ones.

The last category E12 contains those incorrect responses which do not fit in any of the other error categories, including responses that were illegible or could not be understood. The percentages of students who made E12 errors for each item were generally less than 10%. This large summary percentage (77.2%) indicates that a lot of the students made E12 errors but in different items.

Conclusions and implications

The following main conclusions can be drawn: (a) generally the students had the basic ability to solve problems in statistical graphs, and they did better in graph reading and construction tasks than in graph interpretation and evaluation tasks; (b) the difficulty level of the items related to the type of graphs used and the type of questions asked; (c) in the view of statistical significance, the male students did better in graph reading, the female students did better in graph construction, and no gender differences were found in graph interpretation and evaluation; (d) stream and grade level had main effects on students' performance in graph reading, but an interaction effect in graph interpretation, construction, and evaluation; (e) twelve error categories were identified. Most graph convention errors (E4, E5, and E7) and errors related to the use of given information (E10) were more serious than the other error categories.

The findings of this study have implications for teaching and curriculum development. Teachers should assist students to correct graph convention errors, help them develop problem solving abilities in statistical graphs, advice them to properly use their contextual knowledge to solve a graph related problem, and guide them to communicate mathematical ideas clearly. The results of the students' performance in the four aspects of USG and the item difficulty levels identified by the Rasch analysis can be used in curriculum design to be more reflective of and relevant to the actual abilities of students.

Acknowledgement
This study is part of my dissertation study. I would like to thank my supervisor, Dr. Wong Khoon Yoong from National Institute of Education, Nanyang Technological University of Singapore, for his great advice, support and encouragement.

References

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, *18*, 382-393.

Friel, S. N., & Bright, G. W. (1996). *Building a theory of graphicacy: How do students read graphs?* Paper presented at the Annual Meeting of AERA (New York). (ERIC Document Reproduction Service No. ED 395227).

Graham, A. (1989). Statistical investigations in the secondary school. In B. Greer & G. Mulhern (Eds.), *New directions in mathematics education* (149-164). London: Rouuledge.

Hair, J. F., Anderson, R. E., Tatham, R., & Black, W. C. (1995). *Multivariate data analysis with readings* (4th ed.). NJ: Englewood cliffs.

Hawkins, A., Jolliffe, F., & Glickman, L. (1992). *Teaching statistical concepts*. New York: Longman.

Lenton, G., Stevens, B., & Illes, R. (2000). Numeracy in science: Pupils' understanding of graphs. *School Science Review, 82*(299), 15-23.

Mellor, J. M. (1991). *A descriptive study of grade four and grade six students' understanding of bar graphs.* Unpublished master dissertation. Memorial University of Newfoundland.

Ministry of education, Singapore. (2001). *Singapore lower secondary mathematics syllabus*. Retrieved July 29, 2001, from http://www1.moe.edu.sg/syllabuses/doc/Maths_LowSec.pdf.

Mooney, E. S. (1999). *Development of a middle school statistical thinking framework.* Ann Arbor, MI: UMI Dissertation Services.

Padilla, M. J., McKenzie, D. J., & Shaw, E. J. (1986). An examination of the line graphing: Ability of students in grades seven through twelve. *School Science and Mathematics, 86*(1), 20-26.

Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 205-237). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). Amsterdam: IOS press.

Watson, J. M., & Moritz, J. B. (2001). Developing of reasoning associated with pictographs: representing, interpreting, and predicting. *Educational Studies in Mathematics, 48*, 47-81.

Watson, J. M., & Pereira-Mendoza, L. (1996). Reading and predicting from bar graphs. *Australia Journal of language and literacy, 19*(3), 244-258.