

## Evaluating Students' Correlation Graphing Capability Using SOLO Taxonomy

Ken W. Li<sup>1</sup> and Merrilyn Goos<sup>2</sup>

<sup>1</sup>Department of Information and Communications Technology, Hong Kong Institute of Vocational Education (Tsing Yi), HKSAR, CHINA

<sup>2</sup>School of Education, The University of Queensland, Brisbane, AUSTRALIA

<sup>1</sup>Corresponding author: Ken W. Li, e-mail: [kenli@vtc.edu.hk](mailto:kenli@vtc.edu.hk)

### Abstract

Many students often fail to accomplish correlation graphing tasks beyond scatterplot construction. This hinders students' capability of performing subsequent regression modelling tasks. Hence, a qualitative analysis of students' correlation graphing capability should be performed in order to identify which parts of correlation graphing they cannot fully grasp so as to be reinforced. To perform the analysis, an assessment instrument, the SOLO taxonomy of correlation graphing capability was derived from the SOLO taxonomy of Biggs and Collis (1982) consisting of five levels of achievement: Prestructural, Unistructural, Multistructural, Relational and Extended Abstract.

A random sample of twenty-three students studying in tertiary level was drawn to attempt seven questions on an individual basis in a test that was conducted in a computing laboratory. Questions 1-2, 3-5 and 6-7 were used to evaluate how much students understand the given data regarding relations between variables, statistical relations between variables and functional relations between variables respectively. After the SOLO analysis has been performed, the findings reveal information about which tasks students cannot accomplish and why they cannot accomplish. These findings therefore should be able to inform teachers especially novices to think how to structure the teaching and learning activities and enhance students' understanding of specific areas in correlation graphing they cannot fully grasp.

Keywords: scatterplots, regression modelling, levels of achievement

### 1. Introduction

Many students who know how to draw scatterplots have difficulty gaining insights and seeing hidden relationships of data (Chick, 2000). This hinders students' capability of performing subsequent regression modelling tasks. Thus, teachers should be aware of obstacles to the development of graphical understanding of scatterplots in students, otherwise it is difficult to reinforce students' statistical graphing capability. Statistical graphing refers to graph construction, graph characterisation as well as graph inference (Cook & Weisberg, 1997; Li & Goos, 2011).

### 2. Literature Review

An assessment of learning outcomes can be more than assessing how students learn but also provides information to teachers about how to improve pedagogy to support learning (Chance, 1997). Thus, assessment should not focus on checking how well students perform statistical computations and/or how well they construct statistical graphs and charts but on their ability to reason about data; reason about results; and reason about conclusions. The reasoning ability so assessed is associated with the model of statistical thinking developed by Bishop and Talbot (2000). Such assessment generally provides information to both teachers and students how well students understand beyond the statistical procedures and computations they have

used (Gal & Garfield, 1997). As such, assessment frameworks, for example, Bude's hierarchical model of assessment (2006), and SOLO (Structure of the Observed Learning Outcomes) taxonomy of Biggs and Collis (1982) can be used.

Bude (2006) assessed students' statistical understanding using three levels: elementary, intermediate and highest achievement. Elementary level evaluates general understanding of statistical definitions and procedures. Intermediate level requires a deeper understanding of statistical data as well as statistical methods. Highest level refers to the skills of justifying and interpreting statistical results.

The SOLO taxonomy can also be used as a framework to assess how well students accomplish learning tasks. The five levels of achievement they can attain are: Prestructural, Unistructural, Multistructural, Relational and Extended Abstract. Prestructural responses are displayed by students who can attempt simple tasks but they cannot accomplish them. Those students who use one relevant aspect have achieved a unistructural level of achievement. Students who use several aspects but treat them unrelated or unconnected, attain a multistructural level of achievement. Relational level of achievement refers to integrating the relationship between different aspects. In attaining the extended abstract level of achievement, students should be able to deduce relationships.

Although Bude's assessment framework (2006) is closely related to the field of statistics, it does not provide exhaustive assessment as in the SOLO taxonomy of Biggs and Collis (1982). Specifically, Bude's second level achievement which is equivalent to the first four levels of achievement in the SOLO taxonomy does not give clear indications of which parts of statistical methods and understanding of statistical data students do not do well. In addition, Bude (2006) pointed out that an assessment framework of students' statistical ability should be developed according to a specific statistical topic because the skills of reasoning used in different statistical approaches have variation in thought processes. For these reasons, an instrument to assess students' correlation graphing capability was derived from the SOLO taxonomy of Biggs and Collis (1982) in this paper.

### 3. Assessment Instrument

A test was designed to evaluate key aspects of students' statistical thinking and graphing in regression modelling. In the test, a set of real-life data with local context ( $y$  = electricity consumption (terajoules),  $x_1$  = air temperature ( $^{\circ}C$ ),  $x_2$  = relative humidity (%),  $x_3$  = index of industrial production,  $x_4$  = the number of telephone lines,  $x_5$  = composite consumer price index, and  $x_6$  = gas consumption (terajoules)) was given and seven specific questions were designed to evaluate students' responses to each particular task in a preliminary examination of data process. The quantity and scope of data were judged to be within the reach of the students' ability.

Question 1 was used to evaluate how much students understood the given data regarding the data context that was essential for choosing appropriate data in regression modelling. Question 2 was used to check how well students justified the reasonableness and meaningfulness of data measurements. Question 3 was to assess students' knowledge of scatterplot construction and proficiency in using Excel graphing tools. Question 4 focused on an appraisal of students' correlation comprehension. Question 5 appraised students' performance of statistical calculations using Excel. Question 6 checked how well students conducted statistical hypothesis testing and reasoned with testing results. Question 7 aimed at assessing

students' ability to reason with correlation results and deduce its practical implications. According to Bishop and Talbot (2000), the first two questions are equivalent to the task of reasoning about data, the fourth and the sixth questions are similar to the task of reasoning about results, and the last question is consistent with the task of reasoning about conclusions.

A qualitative analysis of students' correlation graphing capability should then be performed in order to identify which parts of correlation graphing they cannot fully grasp so as to be reinforced. To perform the analysis, an assessment instrument, the SOLO taxonomy of correlation graphing capability, was derived from the SOLO taxonomy of Biggs and Collis (1982), and modified in accordance with the cognitive model of graphical comprehension developed by Li and Goos (2011).

The prestructural responses are displayed by students who are able to use an appropriate graphing tool but without utilising graphic features: titles, labels, scales, axis and symbols. Those students who may use one of the graphic features in their scatterplots have achieved a unistructural level of achievement. Students whose scatterplots utilise all the graphic features but treat these as isolated entities and/or unrelated to scattering of data, attain a multistructural level of achievement. Integrating the relationship between the measurement, measurement unit, content and context of data and all the graphic features is regarded as a relational level of achievement. In attaining the extended abstract level of achievement, students should be able to deduce the qualitative relationship between two variables as unrelated, positively related or negatively related and reveal whether or not such relationship matches or mismatches with the empirical phenomena.

#### **4. Research Participants**

A random sample of twenty-three full-time students enrolling in Year 2 of the Higher Diploma in Applied Statistics and Computing (HDASC) course in the Hong Kong Institute of Vocational Education on the Tsing Yi Campus was drawn. This cohort of HDASC students was selected because Regression Modelling is a statistical module taught in their Year 2 study in which the teacher (the first author of this paper) planned for improving classroom teaching practice.

The delivery of the module follows a pattern of 2-hour lectures supported by 1-hour computing laboratory sessions in each of fifteen weeks. The lectures were delivered to the whole class in a lecture theatre equipped with IT equipment and audio-video aids. Each of the three tutorial groups of students was assigned laboratory exercises demanding the analysis, design or implementation of the solutions in a statistical computing laboratory. It was decided that an assessment instrument was designed to evaluate how well the students learn the topic of correlation comprehension.

#### **5. Results**

The first question asked the research (student) participants to hypothesise about possible correlation with pairs of variables based on the data context. Presumably, they were all aware of one common phenomenon in Hong Kong. That is, most households had air conditioners but not heaters and they generally turned on air conditioners in hot weather. For those households who had heaters, they might not turn on their heaters in winter because they found the winter in Hong Kong was not cold enough. Of course, students might say there was no relationship between the electricity consumption and air temperature if they could substantiate their answer by assuming that many households might turn on their heaters in winter. The quality of

students' responses were evaluated according to how well they connected among facts or evidence and deduced the relationship between them if any.

Regarding students' responses to correlation appraisal, the top quality of the response was the one illustrating the underlying relationship between air temperature and electricity consumption, together with adequate grounds, by pointing out that the higher the temperature was, the more electricity would be consumed. The actual response given by students to Q1 ii) is related to an index of industrial production (indicates activity of production in all Hong Kong major industries) in which they pointed out that more machines run or more products produced would lead to a higher electricity consumption. Some other responses, which were reliant on statistical tools, such as correlation calculation or scatterplot construction to deduce the relationship, were of less quality. This was because the level of statistical thinking employed by the students was at the operational rather than strategic level, and means that students did not take the opportunity to cross check whether an underlying phenomenon matched or mismatched with the phenomenon derived from empirical data. These responses with justification were relatively better in quality than those which had no justification.

Question 1 was to evaluate the quality of students' responses to hypothesising about possible correlation with pairs of variables based on the data context. Only 4.3% of students provided a correct answer with grounds based data context. More than one-third (39.1%) of students gave a correct relationship between two variables but did not provide adequate grounds or did not justify the relationship based on data context. 13.1% of students gave a correct relationship but provided incorrect wording sequence. About one-third (34.7%) of students gave a correct relationship between two variables by using statistical graphing or calculation tools. 8.7% of students were unable to assess the relationship between two variables.

Question 2 assessed how well students justified whether the values of given data covered a reasonable and meaningful range with respect to its context, measurement and measurement units. About 22% of students could justify the reasonableness and meaningfulness of data measurement with correct and thorough answers. More than one-third (39.1%) of students gave a correct answer with partial reasons for meaningful range. 13% of students gave a correct answer with justification but it was not specific/irrelevant/not explicit/invalid. 8.7% gave a correct answer but did not give any reasons. 4.3% of students gave a correct answer with partial reasons for meaningful range and highlighted the data range. 4.3% of students provided correct and valid but incomplete answers. 4.3% of students were unable to answer the question directly but gave some relevant information. 8.7% of students did not attempt the question.

Question 3 was to assess students' knowledge of scatterplot construction and proficiency in using Excel graphing tools. About 46% of students demonstrated their good knowledge of correlation graphing and proficiency in using Excel graphing tools. The remaining 54% of students made at least one of these technical mistakes. Improper graph orientation exchanged an independent variable ( $x$ ) and a dependent ( $y$ ) variable so that graph readers or users got confused and subsequently misconceived of the data relationship, that is,  $x$  became a function of  $y$ . Inappropriate graph scales distorted the pattern on a scatterplot and consequently led them to mis-appraise correlation from a scatterplot (e.g., Cleveland et al., 1982). An omission of axis labels misled students to treat graphic features as isolated entities and/or unrelated to correlation pattern. An omission of measurement units concealed the physical meanings and magnitude of data.

Question 4 focused on an appraisal of students' correlation comprehension. About 18% of students could comprehend correlation patterns in scatterplots with valid reasons. 43.2% of students gave incorrect or imprecise answers to this question. Their incorrect answers were due to inappropriate graph scales; or wrong or conflicting reasons. They had given imprecise answers as they provided inexplicit explanations or reasons irrelevant to data scattering. Only 4.3% were unable to estimate the correlation coefficient, and 34.8% of students did not attempt this question.

Question 5 appraised students' performance of statistical calculations using Excel. About 61% of students used Excel tools to accomplish correlation calculation tasks including proper selection and use of correlation function or correlation analysis tool and correct input of data and output of correlation results. 8.7% of students used correct tool and syntax to compute a correlation coefficient but did not interpret Excel results. Approximately 31% of students' Excel proficiency could not be assessed because their computer files were corrupt or unavailable.

Students' responses to Question 6 were evaluated based on two criteria. The first criterion dealt with students' knowledge of Excel syntax and programming skills and the second with their performance of statistical hypothesis testing. A little over one-half of students (52.2%) programmed Excel properly for statistical hypothesis testing. However, it was not possible to assess Excel programming for 43.5% of students because computer files were corrupt or unavailable. In addition, only one of students had used incorrect Excel syntax or programmed Excel incorrectly. For example, a parenthesis was misplaced in the Excel function or the number of paired data ( $n$ ) was mis-counted and varying data count was encountered.

Students' responses to Question 6 were then evaluated to compare how well they performed statistical hypothesis testing. It was found that 39.1% of students accomplished statistical hypothesis testing tasks in which they provided proper formulation of null and alternative hypotheses; correct statistical evidence and decision; sound reasoning with statistical evidence from Excel output as well as statistical implications. About 61% of students failed to complete statistical hypothesis testing tasks. Their failures were due to no/incorrect implications for correlation test results; no/incorrect rejection region; no statistical decisions made; or wrong statistical tools or tests used. Obviously, students did not give the correct rejection region owing to using an incorrect probability distribution; misreading the  $z$ -value (standard normal deviate) from the Excel statistical function; mixing up the rationales of one-sided and two-sided tests, particularly without stating null and alternative hypotheses; or wrong Excel programming. Inappropriate statistical tests or wrong statistical decisions resulted from these technical mistakes and eventually led to drawing an inconsistent conclusion or a wrong implication.

Question 7 aimed at assessing students' ability to reason with correlation results and deduce its practical implications. A little over two-thirds (69.6%) of students responded to correlation deduction and synthesis vaguely and their arguments were not linked to the data context. None of students could deduce the data relationship in a practical context. To interpret correlation beyond the superficial level, students needed to peruse the data and understand them contextually, being regarded as a means of judging the potentiality of variables for proposing a regression model. In dealing with synthesis and deduction, a translation of statistical terms was made in the use of lay language in connection with correlation results but only a few of their deduction tasks could fulfil this general translation requirement.

Questions 3, 4, 6 and 7 formed the basis of evaluating students' overall responses in preliminary examination of data using the SOLO taxonomy, focusing on graph construction, graph characterisation and graph inference. 0.0% (none) of students gave pre-structural responses, but 54.5% of students gave unistructural responses, illustrating that they could construct scatterplot between the measurement, measurement unit, content and context of data. About 41% of students' responses displayed multistructural features in terms of graph construction and graph characterisation. It appears that students could identify and utilise all the graphic features to construct scatterplots (i.e., multistructural) but might not fully integrate the relationship between the measurement, measurement units, content and context of data and all the graphic features (i.e., relational). Only 4.5% of students gave relational responses, illustrating that they could integrate the relationship between the measurement, measurement units, content and context of data. 0.0% (none) of students gave extended abstract responses.

## 6. Conclusion

Students performed better in tasks involving scatterplot construction and correlation calculation, than in tasks involving understanding data; judging data reasonableness, reading scatterplots; performing statistical hypothesis testing; and deducing and synthesising correlation results. In summary, students did better in construction of statistical graphs and statistical calculations than in graph characterisation and inference. This discrepancy between technical (i.e., construction of statistical graphs and statistical calculations) and non-technical (i.e., graph characterisation and inference) examination of data tasks was found. These findings therefore should be able to inform teachers especially novices to think how to structure the teaching and learning activities that can enhance the reasoning skills of students so as to interpret correlation results and deduce its practical implications.

## References

- Biggs, J.B. and Collis, K.F. (1982) *Evaluating the quality of learning: The SOLO taxonomy*, Academic Press, New York.
- Bishop, G. and Talbot, M. (2000) "Statistical thinking for novice researchers in biological science", *Proceedings of the IASE Round Table Conference on Training Researchers in the Use of Statistics*.
- Bude, L. (2006) "Assessing students' understanding of statistics", *Proceedings of the 7<sup>th</sup> International Conference on Teaching Statistics*.
- Chance, B.L. (1997) "Experiences with authentic assessment techniques in an introductory statistics course," *Journal of Statistics Education*, 5(3).
- Chick, H. (2000) "Young adults making sense of data", *Proceedings of the 23<sup>rd</sup> Annual Conference of the Mathematics Education Research Group of Australasia*.
- Cleveland, W.S., Diaconis, P. and McGill, R. (1982) "Variables on scatterplots look more highly correlated when the scales are increased", *Science*, 26, 1138-1141.
- Cook, R.D. and Weisberg, S. (1997) "Graphs in statistical analysis: Is the medium the message?" *Proceedings of the Section on Statistical Education of American Statistical Association*.
- Gal, I. and Garfield, J. (Eds.) (1997) *The assessment challenge in statistics education*, IOS Press, Amsterdam.
- Li, K. and Goos, M. (2011) "Reinforcing students' correlation comprehension", *International Journal of Learning*, 17(1), 261-274.