Analysis of Students' Spatial Skills between Genders

LI Ken W.
Retired Scholar

Introduction

Students' spatial skills affect mathematics achievement (Tarte, 1990). Outhred and Shaw (1999) echoed this statement by spelling out: students who are capable of using graphical methods tend to be more successful in solving non-routine statistical problems. Unfortunately, the definition of spatial skills so far given is broad (Tarte, 1990). But the skills here refer to the ability of constructing scatterplots; reading data patterns in the scatterplots; and ultimately translating the patterns into numerical and verbal forms. The verbal forms make demand on justifying whether data are suitable for correlation analysis; and reasoning with correlation results and deducing its practical implications.

To enhance students' graphical understanding and spatial ability, the cognitive model of correlation comprehension developed by Li and Goos (2011) would be useful. The model starts with pattern recognition processes, interpretive processes then follow and integrative processes accomplish the remaining tasks. Pattern recognition process commences by checking the data encoded on a scatterplot. Interpretive processes are perceptual processes that operate on those patterns to retrieve or construct qualitative and quantitative meanings. Integrative processes are conceptual processes that relate the meanings to the graphic features, such as titles, labels and scales or plotting symbols in a graph.

Research studies reported that students' beliefs about mathematics learning differ between genders. Such beliefs may affect mathematics learning as well as achievement. Girls excel boys on mathematical operations and spatial skills (e.g., Tarte, 1990). This is probably owing to a difference between male and female students' personality characteristics, thus influencing their learning attitudes in various ways (Burr, 1998). For instance, girls are more social and vocal; they prefer the environment that fosters collaborative learning and social interaction more than boys. Female students were more favourable to the use of computers for learning by gaining more from collaborative learning than from individual learning because they were more willing to share ideas; exchange views; and offer mutual assistance, thus resulting in achievement through concerted effort (Fisher, 1993). In general, girls prefer to work on open-ended problems (Hawkins et al., 2014).

A research question naturally arises in the present context is, how students' spatial skills differ between genders? It follows that a qualitative analysis of students' correlation graphing capability should be performed in order to identify which parts of correlation graphing males did better than females, or vice versa.

Review of Literature

The qualitative analysis does not aim at checking how well students perform statistical computations and/or how well they construct statistical graphs and charts, but rather 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 on 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 while 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 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 in accordance with 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.

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 (°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. The quantity and scope of data were judged to be within the reach of the students' ability.

Seven specific questions were designed to evaluate students' responses to each particular task in a preliminary examination of data process. 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. In addition, the third and the fifth question are to assess students' knowledge of scatterplot construction and statistical calculations proficiency in using Excel tools.

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 as developed by Li and Goos (2011). A comparison of the graphing capability between male and female students was made, assuming that both genders were offered equal educational opportunity; they were treated equally in classroom; they achieved equal learning outcomes.

The prestructural responses are displayed by students who are able to use an appropriate graphing tool but without utilizing 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 utilize 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, as well as 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.

Table 1. The SOLO taxonomy of correlation graphing capability

SOLO score	SOLO description	Capacity
1	Prestructural	Construct scatterplots and identify scattering of data without using graphic features: titles, labels, scales, axis, and symbols.
2	Unistructural	Construct scatterplots together with one of these graphic features: titles, labels, scales, axis, and symbols.
3	Multistructural	Spot graphic features but treat as isolated entities and/or unrelated to scattering of data.
4	Relational	Integrate the relationship between the measurement, measurement unit, content and context of data.
5	Extended Abstract	Deduce correlation between two variables.

The test was validated by a professor who has extensive teaching and research experience in the field of mathematics education and assessment. She also oversaw and validated the author's adaptation of the SOLO taxonomy and checked its applicability for analyzing student responses to the test questions.

Research Participants

A random sample of ten female and thirteen male students enrolling in a higher diploma course offered by an academic institution in Hong Kong was drawn. Upon completion of the course, they are eligible for job posts, such as statistical officer and marketing research assistant. The students were selected because they were taught Regression Modelling module incorporating the cognitive model of correlation comprehension. It was decided to evaluate how well the students learn the topic of correlation comprehension using the assessment instrument.

Analysis of Data

Apart from the qualitative analysis of students' responses to the seven questions, it would be better supplemented with a summative answer of whether male or female students outperformed their counterparts by conducting statistical analysis of the responses. Statistical tests were performed to compare the proportions of correct and complete response given to each of the questions between student genders. In addition, the qualitative analysis ends up with a SOLO score summarizing overall achievement of each individual student. Two mean SOLO scores were obtained from two groups of students, male and female. These two scores were compared directly and a t-test was then conducted to reveal whether there is gender difference in SOLO achievement. All the statistical tests conducted were under the assumption of data normality and the interpretation of the test results were made at the 5% level of significance.

Results and Discussions

Question 1 was to evaluate the quality of students' responses to hypothesizing about possible correlation with pairs of variables based on the data context. The quality of their responses were evaluated according to how well they connected among facts or evidence and deduced the relationship between them, if any. A little less than one-third (30.8%) of male students could report correct relationship between two variables with grounds based on data context, compared with one-fifth (20%) of female students although the grounds mostly given by students were inadequate. The top quality of the response was the one illustrating the underlying relationship between two variables, together with adequate grounds. About 31% and 30% of male and female students gave a correct relationship between two variables, but did not justify the relationship based on data context and/or did not provide incorrect wording sequence. Similarly, 30.8%

and 30.0% of male and female students gave a correct relationship between two variables by using statistical graphing or calculation tools without any justification, whereas 7.7% of male students did. Moreover, one-fifth (20%) of female students was unable to assess the relationship between two variables, yet all male students had attempted the question. The 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 it 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. However, these responses with justification were relatively better in quality than those without any justification.

Table 2. Students' responses to correlation appraisal (Question 1)

Response Categories	Frequency	
	Female n=10	Male n=13
The answer was correct and provided grounds based on data context.	0.0%	7.7%
Correct relationship was reported but providing inadequate grounds based on data context.	20.0%	23.1%
Correct relationship was reported but without justifying relationship based on data context.	10.0%	23.1%
Correct relationship was reported but providing incorrect wording sequence and without justifying relationship based on data context.	20.0%	7.7%
The answer was reliant on graphing/calculation and with justifying relationship based on data context.	0.0%	7.7%

The answer was reliant on graphing/calculation but without justifying relationship based on data context.	30.0%	30.8%
Unable to assess the possible relationship between two variables	20.0%	0.0%
Overall	100.0%	100.1%†

Note.

† *Owing to rounding, there may be a slight discrepancy between the overall percentages and 100% as shown in the above table.*

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. Most (90%) female students could gave correct answers, as compared with 84.7% of male students although their answers might not be complete. Specifically, 30% of female students could justify the reasonableness and meaningfulness of data measurement with correct and thorough answers, compared with 15.4% of male students. Similarly, 40% and 38.5% of female and male students gave a correct answer with partial reasons for meaningful range respectively. About 31% of male students gave a correct answer with partial reasons for meaningful range and highlighted the data range or with justification not specific/irrelevant/inexplicit/invalid, whereas 20% of female students gave a correct answer without providing any justifications or reasons at all. In addition, 10% and 15.4% of female and male students did not attempt the question or could not address the question directly.

Table 3. Students' judgement of data reasonableness and meaningfulness (Question 2)

Response Categories	Frequency	
	Female n=10	Male n=13
Gave correct and thorough answer.	30.0%	15.4%
Gave correct answer with partial reasons for meaningful range (non-negative/positive/no extreme values) but nothing for reasonable range.	40.0%	38.5%
Gave correct answer with partial reasons for meaningful range (positive) and highlighted data range.	0.0%	7.7%
Gave correct answer with justification but not specific/irrelevant/not explicit/invalid.	0.0%	23.1%
Gave correct answer without any reasons.	20.0%	0.0%
Unable to answer the question directly but giving some relevant information	0.0%	7.7%
Unattempted	10.0%	7.7%
Overall	100.0%	100.1%†

Note.

† *Owing to rounding, there may be a slight discrepancy between the overall percentages and 100% as shown in the above table.*

Seemingly, lower proportion (15.4% versus 30.0%) of male than female students might have better judgement of data reasonableness and meaningfulness but these figures were not statistically significant ($z = -0.8299$, $p = 0.7967$), implying that no difference between student genders.

Question 3 was to assess students' knowledge of scatterplot construction and proficiency in using Excel graphing tools. About 77.8% and 23.1% of female and male students demonstrated their good knowledge of correlation graphing and proficiency in using Excel graphing tools. None (0.0%) of female students as well as 46.2% of male students could use Excel graphing tools and syntax correctly, but the measurement units or the axis labels were missing. 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. Furthermore, 11.1% and 23.1% of female and male students made at least one of these technical mistakes. Improper graph orientation exchanged an independent variable (x) and a dependent variable (y) 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 to mis-appraise correlation from a scatterplot (e.g., Cleveland et al., 1982).

Table 4. Students' achievement in construction of scatterplot (Question 3)

Response Categories	Frequency	
	Female n=9*	Male n=13
Excel graphing tools and syntax were correctly used.	77.8%	23.1%
Excel graphing tools and syntax were correctly used but the measurement units/axis labels were missing.	0.0%	46.2%
Excel graphing tools were correctly used but improper graph orientation.	11.1%	7.7%

Excel graphing tools were correctly used but at least two of these mistakes: improper graph orientation, inappropriate graph scales, distracting graph background and omission of measurement units/axis labels were found.	11.1%	23.1%
Overall	100.0%	100.1%†

Note.

† *Owing to rounding, there may be a slight discrepancy between the overall percentages and 100% as shown in the above table.*

* *One student was excluded from this analysis of SOLO scores because her computer file was corrupt.*

Apparently, a lower proportion (23.1% versus 77.8%) of male than female students could do scatterplot graphing and had proficiency in using Excel graphing tools. These figures were highly statistically significant ($z = -3.0177$, $p = 0.0025$), implying that female students outperformed male students.

Question 4 focused on an appraisal of students' correlation comprehension. About 23% and 10% of male and female students could comprehend correlation patterns in scatterplots with valid reasons respectively. About 39% and 50% of male and female students gave incomplete, incorrect or imprecise answers to this question respectively. They did not give complete answers as not providing any reasons or justification to substantiate their answers. Their incorrect answers were due to inappropriate graph scales; or wrong reasons. They had given imprecise answers as they provided inexplicit explanations or reasons irrelevant to data scattering. Only 10% of female students could not estimate the correlation coefficient but none was found in male students. Noticeably, 38.5% and 30% of male and female students did not attempt this question respectively.

Table 5. Students' responses to reading scatterplot (Question 4)

Response Categories	Frequency	
	Female n=10	Male n=13
Correct answers and valid reasons were given.	10.0%	23.1%
Correct answers were given but without any reasons.	0.0%	7.7%
Correct direction of data relationship was given but with wrong/no strength and without reasons.	20.0%	7.7%
The strength of data relationship was not given explicitly but a brief description of data scattering.	10.0%	0.0%
Correct answers for strength with reasons were given but no answers for the direction.	10.0%	7.7%
Incorrect answers for strength were given while the direction were not given.	10.0%	0.0%
The answers were incorrect due to inappropriate graph scales.	0.0%	7.7%
The answers were imprecise and without giving any reasons.	0.0%	7.7%
Unable to estimate the correlation coefficient	10.0%	0.0%
Unattempted	30.0%	38.5%
Overall	100.0%	100.1%†

Note.

† Owing to rounding, there may be a slight discrepancy between the overall percentages and 100% as shown in the above table.

It seems that a higher proportion (23.1% versus 10.0%) of male than female students could correctly comprehend correlation patterns in scatterplots. These figures were not statistically significant ($z = 0.8688, p = 08075$), implying that male students could not do correlation comprehension better than female students. Nevertheless, male students gave more varieties of incomplete or incorrect answers.

Question 5 appraised students' performance of statistical calculations using Excel. Male students (61.5%) performed correlation calculation tasks as good as female students (60%). Specifically, they were able to input correlation data correctly; and to select and use correlation analysis tool properly. About 15% of male students used correct tool and syntax to compute a correlation coefficient, but did not interpret Excel results. None was found in female students. Approximately 8% male students' Excel proficiency could not be assessed because their computer files were corrupt or unavailable.

Table 6. Students' responses to correlation calculation (Question 5)

Responses	Frequency	
	Female n=10	Male n=13
The answers as well as Excel tool and syntax were correct.	60.0%	61.5%
Correct Excel tool and syntax were used but no implications were given.	0.0%	15.4%
The answers were correct but unable to assess student's Excel proficiency because computer file was corrupt/unavailable.	40.0%	15.4%
Unable to assess student's Excel proficiency because computer file was corrupt/unavailable.	0.0%	7.7%
Total	100.0%	100.0%

A slightly higher proportion (61.5% versus 60.0%) of male than female students could correctly perform correlation calculation tasks. However, no statistical evidence ($z = -0.0749, p = 0.5298$) shows that male students could do the tasks better than female students.

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. About one-half of female and male students (50% versus 53.8%) programmed Excel properly for statistical hypothesis testing. However, it was not possible to assess Excel programming for 40% and 46.2% of female and male students because computer files were corrupt or unavailable. In addition, only one female student 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. No significant difference in the knowledge of Excel syntax and programming skills between male and female students was found, but it is worth noting that about 40% of students' computer files were corrupt or unavailable irrespective of student genders.

Table 7. Students' knowledge of Excel programming and syntax (Question 6)

Responses	Frequency	
	Female n=10	Male n=13
Correct Excel programming was used.	50.0%	53.8%
Incorrect Excel syntax/programming was used.	10.0%	0.0%
Unable to assess Excel programming because computer file was corrupt/unavailable.	40.0%	46.2%
Correct Excel programming was used.	100.0%	100.0%
Overall	50.0%	53.8%

A slightly higher proportion (53.8% versus 50.0%) of male than female students could have better knowledge of Excel syntax and programming skills but statistical evidence ($z = 0.1831, p = 0.5726$) does not support this finding.

Students' responses to Question 6 were then evaluated to compare how well they performed statistical hypothesis testing. It was found that 40% and 38.5% of female and male 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 60% of female and male 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. Alarmingly, in either gender of student, 30% of students did not specify rejection region or gave wrong rejection region. They 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. Obviously, more male than female responses (i.e., 23.1% versus 10.0%) to this question displayed wrong statistical tools or tests. Wrong statistical decisions that could be resulted from these technical mistakes eventually led to drawing an inconsistent conclusion or a wrong implication.

Table 8. Students' performance of statistical hypothesis testing (Question 6)

Responses	Frequency	
	Female n=10	Male n=13
Correct and complete answers were given.	40.0%	38.5%
The answers were correct but giving no/inadequate statistical evidence.	10.0%	7.7%

No/incorrect rejection region was given.	30.0%	30.8%
No statistical decision was made.	10.0%	0.0%
Wrong statistical tools/tests were used.	10.0%	23.1%
Total	100.0%	100.1%†

Note.

† *Owing to rounding, there may be a slight discrepancy between the overall percentages and 100% as shown in the above table.*

A slightly lower proportion (38.5% versus 40.0%) of male than female students could perform statistical hypothesis testing better, but statistical evidence ($z = -0.0749, p = 0.472$) does not substantiate this claim.

Question 7 aimed at assessing students' ability to reason with correlation results and deduce its practical implications. None of the 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 fulfill this general translation requirement. Male students slightly outperformed female students albeit their vague responses to correlation deduction and synthesis, and moreover their arguments were not linked to the data context irrespective of student genders. It is worth noting that a higher percentage of female students did not attempt the question.

Table 9. Students' correlation deduction and synthesis (Question 7)

Responses	Frequency	
	Female n=10	Male n=13
Correct strength and/or direction of data relationship were given but without deducing the relationship in practical context.	60.0%	76.9%
Unable to synthesize data relationship from Q6 but matched with results from Q1.	10.0%	15.4%
Unrelated matters were highlighted.	10.0%	0.0%
Unattempted	20.0%	7.7%
Total	100.0%	100.0%

It seems that a higher proportion (76.9% versus 60.0%) of male students than female students could correctly assess strength and direction of data relationship but being unable to deduce the linear relationship in a practical context. These figures were not statistically significant ($z = 0.8721, p = 0.8084$), implying no difference between student genders.

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. None (0.0%) of the students gave pre-structural responses or extended abstract responses irrespective of student genders. About 22% and 77% of female and male students gave unistructural responses respectively, illustrating that they could construct scatterplot between the measurement, measurement unit, content, and context of data. Approximately 67% and 23% of female and male students' responses displayed multistructural features respectively in terms of graph construction and graph characterisation, implying that they could identify and utilise all the graphic features to construct scatterplots. Only 11% and 0% of female and male students gave relational responses respectively, illustrating that they could integrate the relationship between the measurement, measurement units, content, and context of data.

Table 10. Frequency distribution of students' SOLO scores of preliminary examination of data

SOLO score	SOLO descriptions	Frequency	
		Female n=9[b]	Male n=13
1	Prestructural	0.0%	0.0%
2	Unistructural	22.2%	76.9%
3	Multistructural	66.7%	23.1%
4	Relational	11.1%	0.0%
5	Extended Abstract	0.0%	0.0%
Mean[a]		2.89	2.23
SD[a]		0.57	0.42

Note.

[a] *A 5-point scale based on student's responses to correlation graphing.*

[b] *One student was excluded from this analysis of SOLO scores because of her corrupt computer file.*

A higher proportion (66.7% versus 23.1%) of female students could attain higher level (i.e., multistructural) of SOLO achievement than their male counterparts. In addition, female students achieved higher SOLO mean scores ($M = 2.89$, $SD = 0.57$) than the male students ($M = 2.23$, $SD = 0.42$). Moreover, these SOLO mean scores were statistically significant, $t(21) = 2.9774, p = 0.0074$, implying that female students attained a higher level of spatial skills than their male counterparts.

Summary

Female SOLO performance on the average was significantly higher than that of their male counterparts, illustrating better knowledge of scatterplot construction and proficiency in using Excel graphing tools. Nevertheless, both male and female students did poorly in the tasks of reasoning correlation results and deducing its practical implications. Noticeably, low-spatial students were not hampered by their low level of spatial skill as they could make use of alternative skill like computing to accomplish correlation graphing tasks.

All statistical conclusions, which were drawn at the five percent level of significance, could only indicate a high likelihood of statistical significance or non-significance, rather than generalizing definite answers. All statistical tests were conducted under the assumption of data normality. Apart from the statistical assumption, the comparison of the graphing capability between male and female students so far was made on the assumption of gender of students, the fair assessment of learning outcomes, equal learning opportunity being offered by the teacher and being treated equally by their classmates or teacher in classroom.

References

Biggs, J.B. and Collis, K.F. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* New York: Academic Press.

Bishop, G. and Talbot, M. (2000). Statistical thinking for novice researchers in biological science. In C. Batanero (Ed.), *Proceedings of the IASE Round Table Conference on Training Researchers in the Use of Statistics*, 215-226. Tokyo: International Association for Statistical Education.

Bude, L. (2006). Assessing students' understanding of statistics. In A. Rossman & B. Chance (Eds.), *Proceedings of the 7th International Conference on Teaching Statistics.* Brazil: International Association for Statistical Education.

Burr, V. (1998). *Gender and social psychology.* London: Routledge Press.

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.

Fennema, E. (1990). Teachers' beliefs and gender differences in Mathematics. In E. Fennema & F.C. Leder (Eds.), *Mathematics and Gender*, 169-187. New York: Teachers College Press.

Fisher, E. (1993). The teacher’s role. In P. Scrimshaw (Ed.), *Language, classrooms and computers*, 51-74. London: Routledge.

Fisher, E. (1997). Educationally important types of children's talk. In D.J. Corson (Series Ed.) & Wegerif & P. Scrimshaw (Vol. Eds.), Language and education library: vol. 12. Computers and talk in the primary Classroom, 22-37. Clevedon: Multilingual Matters Ltd.

Gal, I. and Garfield, J. (Eds.) (1997). *The assessment challenge in statistics education.* Amsterdam: IOS Press.

Hawkins, A., Jolliffe, F., and Glickman, L. (2014). *Teaching statistical concepts.* Oxford: Routledge.

Li, K. and Goos, M. (2011). Reinforcing students’ correlation comprehension. *International Journal of Learning, 17(1)*, 261-274.

Outhred, L. and Shaw, P.F. (1999). Visual representations in first statistics. In J.M. Truran & K.M. Truran (Eds.), *Proceedings of the 22nd Annual Conference of the Mathematics Education Research Group of Australasia*, 411-418. Adelaide: Mathematics Education Research Group of Australasia.

Tarte, L. (1990). A. Spatial skills, gender and mathematics. In E. Fennema and G.C. Leder (Eds.), *Mathematics and Gender*, 27-59. New York: Teachers College Press.

Author’s e-mail: dr_ken_li@yahoo.com.hk