International Journal of Innovation, Management and Technology, Vol. 1, No. 1, April 2010
ISSN: 2010-0248

Measuring Students' Understanding of Statistical Concepts using Rasch Measurement

Noor Lide Abu Kassim, Nor Zatul-Iffa Ismail, Zamalia Mahmud, and Mohammad Said Zainol

***Abstract* — This study investigates students' conceptual knowledge and understanding of basic statistical concepts and compares it against statistical competence, which is associated with discrete statistical knowledge and basic interpretive skills. It particularly examines the correspondence between students' perceived ability and their empirical understanding of the concepts. Two instruments were developed: a 20-item test to measure students' empirical understanding of the basic statistical concepts and a questionnaire with matching items to measure their perceived ability of these concepts. For a direct comparison of the two, students' responses to the test and questionnaire items were jointly analyzed using the Rasch measurement model. Results of the analysis show that conceptual understanding of basic statistical concepts is more difficult to attain than statistical competence. The results also suggest that students more often than not overestimated their understanding of basic statistical concepts, particularly those requiring conceptual understanding of the concepts.**

***Index Terms*—basic statistical concepts, conceptual understanding, statistical competence, Rasch measurement,**

I. Introduction

In a number of studies, students in statistics courses were found to describe rather than justify their statistical solutions [1] and fail to establish a conceptual base for their solution strategies [2]. Although students may be able to answer some test items correctly or perform calculations correctly, they may still misunderstand basic ideas and concepts in statistics. This may be explained by the lack of conceptual understanding of what is being constructed or how statistical concepts are interrelated. For example, Garfield et al. [3] and Bakker [4] found in their studies that although students may be able to calculate basic statistics, a sound understanding of what was being constructed or how statistical concepts are interrelated is rare. Clark et al. [5] and Matthews et al. [6], similarly, found that students who receive top grades in a class may not understand and remember the basic ideas of statistics.

The lack of conceptual knowledge and understanding of statistical concepts is particularly seen in relation to basic statistical concepts such as reasoning about distributions and graphical representations of distributions [7,8,9], understanding concepts related to statistical variation such as measures of variability [10,11]), and sampling distributions [12,13,14]. This state of affairs is unfortunate given that statistical reasoning is crucial in dealing with the prevalence of statistical data in the media and other sources of information that pervades our daily life.

This study, therefore, seeks to investigate students' conceptual knowledge and understanding of basic statistical concepts and compare it against statistical competence, which is associated with discrete knowledge and basic interpretive skills. This study particularly examines the correspondence between students' empirical understanding and their perceived ability in these basic statistical concepts. As conceptual knowledge and understanding has been reported to be overestimated by students, it would be of interest to investigate where this is most apparent. It is hoped that the results of this study would provide insight to the development of students' conceptual knowledge and understanding that underlies statistical reasoning and thinking.

II. Literature Review

A. *Statistics as a Discipline*

Statistics as a discipline can be described as the mathematical study of the likelihood or probability of events occurring based on known information and inferred by taking a limited number of samples [15]. A dictionary of mathematics terms defined statistics as "the study of ways to analyze data… it consists of descriptive statistics and statistical inference" [16,17] in the same vein defined statistics as "the science of collecting, organizing, analyzing, and interpreting data in order to make decisions". These descriptions imply that statistics as a discipline is used to make sense of data for use in the decision-making process.

For many instructors, statistics is an area of applied mathematics that readily lends itself to real-world applications of mathematics. This applied mathematics incorporates sampling techniques and probability to describe

Manuscript received December 18, 2009. This work was supported in part by the_Research Management Institute under the provision of Universiti Teknologi MARA Excellence Fund.

Noor Lide Abu Kassim is a senior lecturer at the Institute of Education, International Islamic University Malaysia, 53100 Jalan Gombak, Selangor, Malaysia (phone: 6019-2155768; e-mail: noorlide@iium.edu.my).

Nor Zatul-Iffa Ismail is a Ph.D student at the Centre of Statistical Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor , MALAYSIA (e-mail: iffa@tmsk.uitm.edu.my).

Zamalia Mahmud is an Associate Professor at the Centre of Statistical Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor , MALAYSIA (e-mail: zamal669@salam.uitm.edu.my).

Mohammad Said Zainol is a Professor at the Centre of Quantitative Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, MALAYSIA (e-mail: said@tmsk.uitm.edu.my).

IACSIT
International Association of
Computer Science and Information Technology
WWW.IACSIT.ORG

and predict outcomes based on experience. It could be argued that accurate interpretation of statistics is a necessary prerequisite for an informed and educated citizenry today since statistical data, summaries, and inferences appear frequently in the everyday lives of people than any other form of mathematical information [18].

Statistics in all its complexities may create bias and misuse in many different ways [19]. Citizens need to be aware that statistics can be used to manipulate and deceive through misrepresentation, such as by repeating studies until desirable results are obtained, or by using small and/or biased samples. Thus, it is critical for students, the future generation of adult society, to be statistically literate in order to critically analyze the information they receive.

B. Statistical Literacy, Statistical Competence, Conceptual Knowledge, and Statistical Reasoning

What is statistical literacy? A common theme in the literature is that statistical literacy does not only involve knowledge and understanding of statistical concepts such as distribution, probability and sampling but it is also concerned with the ability to critically evaluate the adequacy of concepts that have been applied without proper statistical foundation [20,21]. Students need to reason and make sense of statistical information. Having statistical literacy, therefore, means being able to describe and summarize basic data as well as being able to understand and explain statistically complex concepts such as trends. In other words, it requires the ability to extract, understand, and explain data that is presented in a variety of ways. To be statistically literate one must understand that how data is organized can contribute to how it is interpreted [8,20,22,23].

Conceptual knowledge, on the other hand, has been characterized as knowledge that is rich in relationships, where discrete pieces of statistical knowledge, ideas and concepts are connected to construct a network of interrelated propositions [24,25]. The construction of interrelated propositions can occur "between pieces of information already stored in the memory or between an existing piece of knowledge and one that is newly learned [24]. Conceptual knowledge and understanding, therefore, is essential for the development of statistical reasoning and thinking. Without it, students would not be able to make connections and explain the relationships between the different statistical processes or discrete statistical knowledge [25]. The relationship between conceptual knowledge and understanding and statistical reasoning is clearly explicated in the following definition of statistical reasoning given by Garfield [23]; Statistical reasoning is the way people reason with statistical ideas and make sense of statistical information. This involves making interpretations based on sets of data, graphical representations, and statistical summaries. Much of statistical reasoning combines ideas about data and chance, which leads to making inferences and interpreting statistical results. Underlying this reasoning is a conceptual understanding of important ideas, such as distribution, center, spread, association, uncertainty, randomness, and sampling [23].

Statistical competence is held to include the following components: (i) data awareness, (ii) an understanding of certain basic statistical concepts and terminology, (iii) knowledge of the basics of collecting data and generating descriptive statistics, (iv) basic interpretation skills (the ability to describe what the results mean in the context of the problem), and (v) basic communication skills (being able explain the results to someone else) [26]. This basic knowledge, Rumsey [26] argues, underlies statistical reasoning and thinking. Statistical literacy, conceptual knowledge and understanding, and statistical reasoning, thus, can be thought of as unique areas in themselves and may be represented as a hierarchy, where statistical competence provides a foundation for conceptual knowledge and understanding, and conceptual understanding for statistical reasoning (see Figure 1).

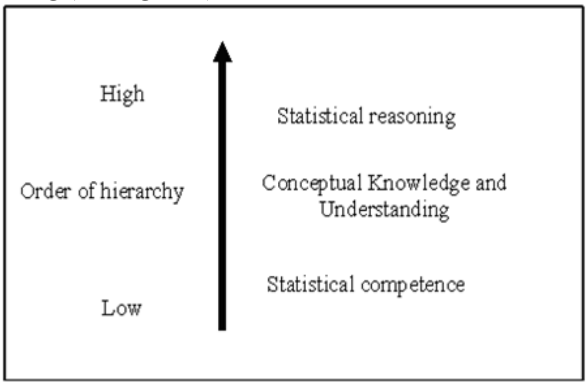


Figure 1. Hierachy of relationship between statistical competence, conceptual knowledge and understanding and statistical reasoning.

C. Teaching and Learning of Statistics

The teaching and learning of statistics that focuses on computation skills and discrete statistical knowledge has come under considerable scrutiny in recent years and a number of recommendations have been made to develop students' conceptual understanding and statistical reasoning. For example, the American Statistical Association, in 2005, endorsed a set of instructional guidelines published in the Guidelines for Assessment and Instruction in Statistics Education (GAISE) project [27]. In the guidelines, the following recommendations for statistics education were made:

- Emphasize statistical literacy and develop statistical thinking;
- Use real data;
- Stress conceptual understanding rather than mere knowledge of procedures;
- Foster active learning in the classroom;
- Use technology for developing conceptual understanding and analyzing data;
- Use assessments to improve and evaluate student learning

Gal et al. [28], in line with the call for reform in the teaching and learning of statistics, outlined eight instructional goals for statistics education to help students understand and use statistical information and data in an increasingly information-dense society. These goals are to have students understand the big ideas that underlie statistical inquiry which include:

1) Understand the big ideas that underlie statistical inquiry. These ideas include:
    - The existence of variation

- The need to describe populations by collecting data
- The need to reduce raw data by noting trends and main features through summaries and displays of the data
- The need to study samples instead of populations and to infer from samples to populations
- The logic behind related sampling processes
- The notion of error in measurement and inference, and the need to find ways to estimate and control errors
- The need to identify causal processes or factors
- The logic behind methods (such as experiments) for determining causal processes

2) Understand the method of statistical investigations. This includes study planning, data planning, data collecting and organizing, data analysis, interpretation of results, conclusions and implications.
3) Become proficient in procedural skill
4) Understand the relationship between the mathematical parts (raw data, graphs, summary statistics, etc) and how changes in data affect these.
5) Understand probability and chance where the emphasis is on an informal grasp of probability and an understanding of the commonly used language
6) Develop interpretive skills and statistical literacy in order to become effective users of statistical information and be able to critically analyze and question it.
7) Develop the ability to communicate well and use statistical and probability terminology.
8) Develop an appreciation for statistical methods as a tool.

III. Method

A. Participants

The research method utilized for this study was a survey method with two data collection instruments: a self-report questionnaire and a test of basic statistical knowledge. The self-report questionnaire and the test were administered separately to ensure that neither one influences the response to the other. The questionnaire was administered first before the test. Total time given for completion of the questionnaire and test was 40 minutes.

In this preliminary study, purposive sampling was used. Several introductory statistics courses conducted at public tertiary institutions were identified and the instructors contacted at the onset of this study. However, only two instructors from two institutions agreed to involve their students in the study. Subsequently, a total of 115 students from these two institutions made up the sample for this study.

B. Data Collection Instruments

A 20-item test was developed to measure students' statistical competence and conceptual knowledge and understanding of the basic statistical concepts. The concepts tested focused on types of data, graphical representations of distributions, measures of sampling distribution, and measures of variability. These topics were selected for the following reasons:

1) Studies have shown that students have difficulty with reasoning about distributions and graphical representations of distributions [7,8], understanding concepts related to statistical variation such as measures of variability [10,11] and sampling distributions [12,13]; and
2) Proficiency in statistics is always related to a specific topic. Thus, assessment should not be focused on general competencies, but should be focused on students' knowledge of specific topics and try to gauge their understanding of the subject matter [29].

The questionnaire to measure students' perceived ability was developed alongside the test to allow for comparisons between the two.. Each item has a 5-point Likert-type response format ranging from (1) Strongly Disagree to (5) Strongly Agree.

C. Data Analysis Procedure

The data was analyzed using WINSTEPS, version 3.64.2. In the initial analysis, responses to the test items and the statements in the questionnaire were analyzed separately. A joint analysis was then conducted where the responses to the statements on the questionnaire were collapsed into a dichotomy to complement the response scale of the test. This was done to allow for a more direct comparison between empirical understanding and perceived ability. In the analysis, the following were also examined: (i) the validity of items and student responses, (ii) the capacity in which the items were able to define a continuum of increasing intensity, (iii) reliability, (iv) unidimensionality, and (iv) construct definition.

IV. Results

A. Reliability

From Figure 2, the reliability of item difficulty estimates is high (.96). The item separation index of 4.66 indicates that the items can be separated into 4 difficulty strata. As item reliability indicates the ability of the test to reproduce the hierarchy of items along the measured variable [30,31], a reliability coefficient of .96 suggests that this order of item hierarchy will be replicated with a high degree of probability if the items were given to other comparable cohorts. With regard to person measures, the reliability coefficient is considerably lower at 0.71. This is attributed to the considerable misfitting responses in the data. Responses to the statements in the questionnaire, on the other hand, showed greater consistency and this showed in a higher reliability coefficient for the questionnaire data.

```
+-----------------------------------------------------------------------------+
| PERSONS     115 INPUT     104 MEASURED              INFIT         OUTFIT    |
|           SCORE     COUNT     MEASURE   ERROR     IMNSQ   ZSTD  OMNSQ   ZSTD|
| MEAN      23.7      37.4         .85     .44        .99     .0    .98    -.1|
| S.D.       5.5       4.0         .81     .04        .29    1.5    .48    1.4|
| REAL RMSE    .44  ADJ.SD     .68  SEPARATION  1.55  PERSON RELIABILITY  .71|
|-----------------------------------------------------------------------------|
| ITEMS      40 INPUT      40 MEASURED                INFIT         OUTFIT    |
| MEAN      61.5      97.2         .00     .28        .97     .1    .96     .3|
| S.D.      20.5       8.3        1.35     .05        .14    1.3    .29    1.6|
| REAL RMSE    .28  ADJ.SD    1.32  SEPARATION  4.66  ITEM   RELIABILITY  .96|
+-----------------------------------------------------------------------------+
```

Figure 2. Person and Item Reliability Coefficients

IACSIT
International Association of
Computer Science and Information Technology
WWW.IACSIT.ORG

B. Items and Persons Distributions

One of the most important features of the Rasch approach is that students' scores and item difficulty are transformed onto one scale so that they are related to [32,33]. This allows item difficulty and person ability for a group of examinees on a group of items to be directly compared. This is known as 'mapping', where estimates of person ability and item difficulty are represented graphically in the form of an item-by-person map (see Fig. 3). Since both the items and persons are represented graphically on the same logit scale, it is possible to see if the items fit the ability of the students.

From the map it is evident that a large number of items can be found along the continuum on which the majority of students' abilities fall. However, there are items at the difficult and easy ends where a minimal number of students could be found.

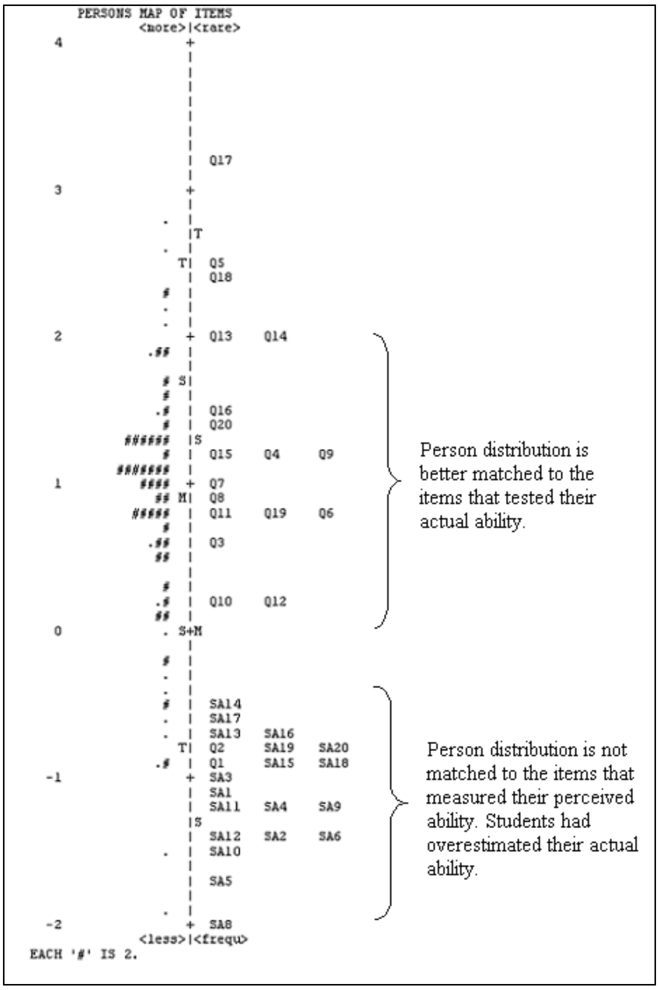


Figure 3. Hierarchy of relationship

Fig. 3 indicates the distribution of items from the test as well as the self-report questionnaire. As expected, items on the questionnaire which were self-reported clustered towards the bottom of the logit scale whereas the test items are more evenly distributed along it. This indicates that the participants had largely overestimated their actual ability. The most difficult item on the test is Item 17. Two items – Item 1 and Item 2 – were the easiest on the test and are clustered together with the self-report (perceived ability) items.

Person distribution is better matched to the items that tested their actual ability compared to items that measured their perceived ability. This indicates that the participants had overestimated their actual ability. Persons are also largely clustered at the middle of the scale (between -1 logit and +2 logits) where most of the items are located. This suggests that the items are not functioning well enough to clearly separate persons into differing levels of ability.

TABLE 1. HIERARCHY OF ITEMS BASED ON 'PERCEIVED ABILITY'

SA8	-2.00	SU	DISTRIBUTION-GRAPHIC PRESENTATION
SA5	-1.71	SC	DATA-GRAPHIC PRESENTATION
SA10	-1.54	SU	MEASURES OF CENTRE
SA6	-1.45	SC	DISTRIBUTION-GRAPHIC PRESENTATION
SA12	-1.44	SC	MEASURES OF CENTRE
SA7	-1.40	SU	DISTRIBUTION-GRAPHIC PRESENTATION
SA2	-1.38	SC	DATA
SA11	-1.18	SU	MEASURES OF CENTRE
SA9	-1.18	SU	DISTRIBUTION-GRAPHIC PRESENTATION
SA4	-1.16	SC	DATA
SA1	-1.06	SC	DATA
SA3	-1.00	SC	DATA
SA15	-0.89	SU	MEASURES OF SPREAD
SA18	-0.89	SU	MEASURES OF SPREAD
SA20	-0.82	SU	MEASURES OF SPREAD
SA19	-0.80	SU	MEASURES OF SPREAD
SA13	-0.72	SU	MEASURES OF CENTRE
SA16	-0.67	SU	MEASURES OF SPREAD
SA17	-0.61	SU	MEASURES OF SPREAD
SA14	-0.46	SU	MEASURES OF SPREAD

TABLE 2. HIERARCHY OF ITEMS BASED ON PERFORMANCE ON TEST

Q1	-0.90	SC	DATA
Q2	-0.76	SC	DATA
Q12	0.20	SC	MEASURES OF CENTRE
Q10	0.20	SU	MEASURES OF CENTRE
Q3	0.61	SC	DATA
Q6	0.79	SC	DISTRIBUTION-GRAPHIC PRESENTATION
Q19	0.79	SU	MEASURES OF SPREAD
Q11	0.83	SU	MEASURES OF CENTRE
Q8	0.92	SU	DISTRIBUTION-GRAPHIC PRESENTATION
Q7	1.01	SU	DISTRIBUTION-GRAPHIC PRESENTATION
Q4	1.18	SC	DATA
Q15	1.22	SU	MEASURES OF SPREAD
Q9	1.22	SU	DISTRIBUTION-GRAPHIC PRESENTATION
Q20	1.45	SU	MEASURES OF SPREAD
Q16	1.49	SU	MEASURES OF SPREAD
Q14	1.99	SU	MEASURES OF SPREAD
Q13	2.04	SU	MEASURES OF CENTRE
Q18	2.40	SU	MEASURES OF SPREAD
Q5	2.46	SC	DATA-GRAPHIC PRESENTATION
Q17	3.21	SU	MEASURES OF SPREAD

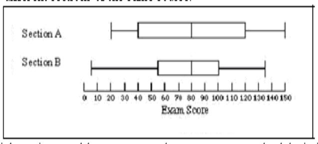
Table 1 and Table 2 present the hierarchy of items based on their difficulty estimates for both the test data and the questionnaire data. From the tables, it is evident that participants have somewhat accurately estimated the relative difficulty of measures of spread and measures of centre. However, they underestimated their knowledge about types of data (with the exception of ordinal data), and overestimated their knowledge of graphic presentation of data. Furthermore, it was also found that students in most cases overestimated their ability in the basic statistical concepts.

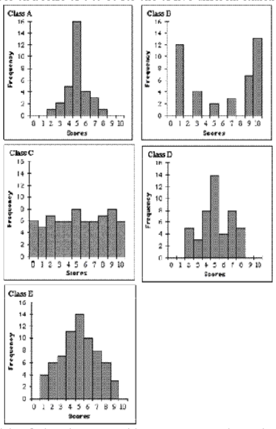
For the purpose of comparing the differences between the items (perceived ability versus performance), we chose to compare several items relating to Types of Data and Graphic Representation of Data concepts. This can be viewed respectively in Table 3 and Table 4.

TABLE 3. PERCEIVED ABILITY AND EMPIRICAL UNDERSTANDING (TYPES OF DATA)

Students' perceived ability to understand concepts	Students' objective understanding of concepts
SA1: I am able to identify the various types of numerical data	Q1. Which of the following is a numerical data? A. number of questions answered correctly B. test type (motivational or threatening) C. whether a particular question was correct or not correct D. whether the letter the instructor wrote was favorable or unfavorable
SA2: I am able to identify the various types of categorical data.	Q2. Which of the following is a categorical data? A. number of questions answered correctly B. test type (motivational or threatening) C. the number of tests distributed D. the number of students in each group
SA3: I am able to identify the various types of ordinal data.	Q3. Which of the following is an ordinal data? A. the rank of the students scores in the class B. test type (motivational or threatening) C. whether a particular question was correct or incorrect D. the number of students in each group

TABLE 4. PERCEIVED ABILITY AND EMPIRICAL UNDERSTANDING (GRAPHIC REPRESENTATION OF DATA)

Students' perceived ability to understand concepts	Students' objective understanding of concepts
SA16: I am able to determine which of two box-plots represents a larger standard deviation.	Q16. The two box-plots below display final exam scores for all students in two different sections of the same course.  Which section would you expect to have a greater standard deviation in exam scores? A. Section A. B. Section B. C. Both sections are about equal. D. It is impossible to tell.

Students' perceived ability to understand concepts	Students' objective understanding of concepts
SA18: I am able to correctly estimate standard deviations for different distributions as displayed by the histograms.	Q18. Five histograms are presented below. Each histogram displays test scores on a scale of 0 to 10 for one of five different statistics classes  Which of the classes would you expect to have the lowest standard deviation, and why? A. Class A, because it has the most values close to the mean. B. Class B, because it has the smallest number of distinct scores. C. Class C, because there is no change in scores. D. Class A and Class D, because they both have the smallest range. E. Class E, because it looks the most normal.

For illustration, estimates for students' perceived ability and empirical understanding relating to types of data and representation of data are extracted and displayed in Table 5.

TABLE 5. ABILITY AND PERFORMANCE ESTIMATES

Types of Data	Graphic Representation of Data
SA1 = -1.06, Q1 = -0.90 SA2 = -1.38, Q2 = -0.76 SA3 = -1.00, Q3 = 0.6	SA16 = -0.67, Q16 = 1.49 SA18 = -0.89, Q18 = 2.40

These differences are clearly depicted in the Item Characteristic Curves (ICCs) where large discrepancies in perceived ability and empirical understanding are evident. In Fig. 4 there exists incongruence between SA1 and Q1, and SA2 and Q2 where the students underestimated their knowledge about types of data particularly in the numerical and categorical data. There exists greater incongruence between SA3 and Q3 where the students overestimated their knowledge about types of data particularly at the ordinal data. In Fig. 5, there exists greater incongruence between SA16 and Q16 and between SA18 and Q18. In this situation, students also overestimated their knowledge about graphic representation of data.

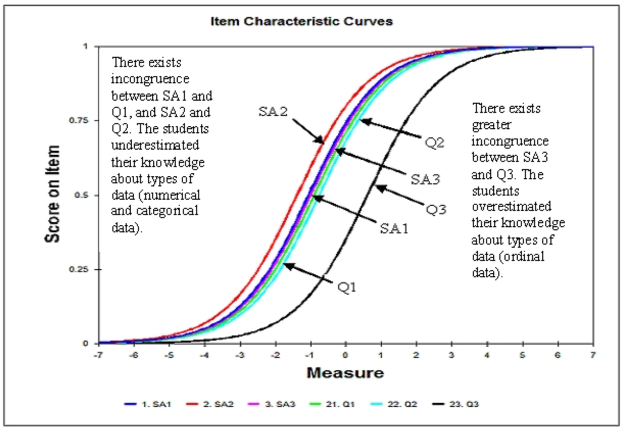


Figure 4. Incongruence between perceived ability and empirical understanding (Types of data)

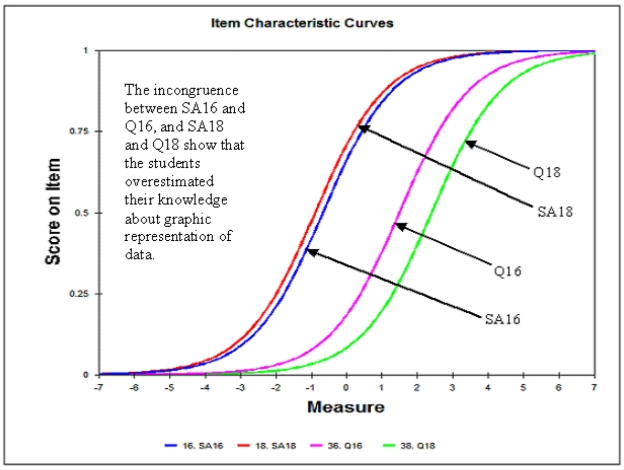


Figure 5. Incongruence between perceived ability and empirical understanding (Graphic Representation of Data)

Table 6 presents the item statistics in correlation order. It can be seen that the point measure correlation for Q18 shows a negative point measure correlation (-0.11). This indicates that the item is problematic and should be considered for improvement in terms of its construction or replacement with a similar item type. The infit statistics are within the cut off range of 0.77 to 1.30. Five items, however, showed fit statistics that are above the 1.30 threshold. This is indicative of outlying responses to the items. However, they are not considered a threat to validity. Further discussion of fit statistics and threats to validity can be seen in Linacre [31].

TABLE 6. CORRELATION ORDER OF ABILITY AND UNDERSTANDING ITEM STATISTICS

ITEM STATISTICS: CORRELATION ORDER

ENTRY NUMBER	RAW SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEA CORR.	EXACT MATCH OBS%	EXACT MATCH EXP%	ITEM	G
38	21	104	2.40	.25	1.21	1.4	1.87	3.2	-.11	79.8	80.1	Q18	2
29	44	104	1.22	.21	1.24	3.5	1.42	3.7	.01	50.0	63.8	Q9	2
30	67	104	.20	.22	1.26	2.8	1.35	2.8	.04	59.6	70.0	Q10	2
28	51	104	.92	.21	1.22	3.4	1.43	4.3	.04	50.0	63.1	Q8	2
27	49	104	1.01	.21	1.21	3.3	1.32	3.2	.07	48.1	63.0	Q7	2
37	11	104	3.21	.33	1.11	.5	1.10	.4	.07	89.4	89.4	Q17	2
32	67	104	.20	.22	1.20	2.2	1.21	1.8	.13	57.7	70.0	Q12	2
36	38	104	1.49	.22	1.12	1.7	1.26	2.0	.14	61.5	67.0	Q16	2
35	44	104	1.22	.21	1.06	1.0	1.36	3.2	.20	69.2	63.8	Q15	2
34	28	104	1.99	.23	1.04	.4	1.08	.5	.23	76.9	74.2	Q14	2
26	54	104	.79	.21	1.07	1.1	1.09	1.0	.26	58.7	63.7	Q6	2
25	20	104	2.46	.26	1.01	.1	.91	-.3	.26	80.8	81.0	Q5	2
5	72	80	-1.71	.39	1.05	.3	.86	-.2	.28	88.8	90.0	SA5	1
39	54	104	.79	.21	1.05	.9	1.05	.6	.28	56.7	63.7	Q19	2
33	27	104	2.04	.23	.96	-.3	1.08	.5	.29	77.9	75.0	Q13	2
23	58	104	.61	.21	1.03	.5	1.10	1.0	.29	62.5	65.2	Q3	2
22	84	104	-.76	.26	1.04	.3	1.00	.1	.30	79.8	81.8	Q2	2
7	79	90	-1.40	.34	.96	-.1	1.09	.4	.30	87.8	88.0	SA7	1
40	39	104	1.45	.21	.96	-.5	1.03	.3	.34	72.1	66.3	Q20	2
24	45	104	1.18	.21	.96	-.7	1.11	1.1	.35	69.2	63.5	Q4	2
31	53	104	.83	.21	.99	-.1	.97	-.2	.36	66.3	63.5	Q11	2
8	85	92	-2.00	.41	.89	-.3	.55	-1.0	.42	92.4	92.4	SA8	1
21	86	104	-.90	.28	.94	-.3	.77	-1.0	.43	81.7	83.4	Q1	2
12	82	93	-1.44	.34	.87	-.5	.73	-.8	.44	90.3	88.4	SA12	1
14	68	89	-.46	.27	.89	-.8	.90	-.5	.46	83.1	78.4	SA14	1
6	81	92	-1.45	.34	.86	-.5	.68	-.9	.46	90.2	88.3	SA6	1
13	76	95	-.72	.27	.89	-.6	.77	-1.1	.47	81.1	81.2	SA13	1
15	68	83	-.89	.30	.85	-.8	.91	-.3	.48	85.5	82.8	SA15	1
20	70	87	-.82	.29	.89	-.6	.78	-.9	.48	81.6	81.5	SA20	1
10	81	91	-1.54	.35	.86	-.5	.54	-1.4	.50	90.1	89.1	SA10	1
1	73	87	-1.06	.31	.83	-.8	.74	-.9	.51	86.2	84.5	SA1	1
4	84	98	-1.16	.31	.83	-.8	.69	-1.1	.51	87.8	86.2	SA4	1
18	78	95	-.89	.28	.81	-1.1	.90	-.3	.52	83.2	82.9	SA18	1
9	59	70	-1.18	.35	.79	-.9	.81	-.5	.53	87.1	84.9	SA9	1
2	78	89	-1.38	.34	.81	-.8	.63	-1.2	.53	89.9	87.9	SA2	1
16	75	95	-.67	.27	.81	-1.3	.83	-.8	.54	83.2	80.3	SA16	1
11	81	95	-1.18	.31	.79	-1.1	.64	-1.4	.56	87.4	85.7	SA11	1
19	73	91	-.80	.28	.80	-1.2	.69	-1.5	.57	83.5	81.3	SA19	1
3	82	98	-1.00	.29	.78	-1.2	.63	-1.5	.58	85.7	84.3	SA3	1
17	76	97	-.61	.26	.78	-1.5	.64	-1.9	.60	81.4	79.8	SA17	1
MEAN	61.5	97.2	.00	.27	.97	.1	.96	.3		76.9	77.7		
S.D.	20.5	8.3	1.35	.06	.14	1.3	.29	1.6		12.8	9.6		

V. CONCLUSION

The results of analysis indicate that congruence between empirical understanding and perceived ability was not evident. The difference in student performance on items measuring conceptual understanding and statistical competence was considerable. The initial results also suggest that students substantially overestimated their conceptual knowledge and understanding of the basic statistical concepts. In terms of instrument refinement several actions will have to be taken. Some of the existing items will have to be reviewed to explain unexpected performance. More items will have to be written and the construct definition of the instruments reviewed to provide a much more accurate and valid measurement of the intended construct.

REFERENCES

[1] C.M. Allwood,, and H. Montgomery. Detection of errors in statistical problem solving. *Scandinavian Journal of Psychology. 1982, 23,* 131-140.

[2] C.M. Allwood, and H. Montgomery. Knowledge and technique in statistical problem solving. *European Journal of Science Education. 1981, 3,* 431-450.

[3] J. Garfield, and D. Ben-Zvi (2004). Research on statistical literacy, reasoning, and thinking: issues, challenges, and implications. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking,* (pp. 397-410). Dordrecht, The Netherlands: Kluwer Academic Publishers.

[4] A. Bakker. Reasoning about shape as a pattern in variability. *Statistics Education Research Journal. 2004, 3*(2), 64-83. [Online]. Available: http://www.stat.auckland.ac.nz/serj]

[5] J. Clark, G. Karuat, D. Mathews, and J. Wimbish, J. *The Fundamental Theorem of Statistics: Classifying Student understanding of basic statistical concepts*, 2003. [Online]. Available: http://www1.hollins.edu/faculty/clarkjm/stat2c.pdf]

[6] D. Mathews, and J. Clark. *Successful Students' Conceptions of Mean, Standard Deviation and the Central Limit Theorem.* Unpublished paper. 2003. [Online]. Available: http://www1.hollins.edu/faculty/clarkjm/stats1.pdf.]

[7] A. Bakker, and K. Gravemeijer. Learning to reason about distribution. In D. Ben-Zvi and J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking. 2004,* (pp. 147-168). Dordrecht, The Netherlands: Kluwer.

[8] D. Ben-Zvi. Reasoning about data analysis. In D. Ben-Zvi & J. B. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking. 2004,* (pp. 121-146). Dordrecht, Netherlands: Kluwer.

[9] J. K. Hammerman, and A. Rubin. Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal. 2004, 3*(2), 17-41.

[10] R. delMas, and Y. Liu. Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal. 2005, 4*(1), 55-82. [Online]. Available: http://www.stat.auckland.ac.nz/~iase/serj/SERJ4(1)_delMas_Liu.pdf]

[11] D. Mathews, and J. Clark. Successful students' conceptions of mean, standard deviation, and the Central Limit Theorem. Paper presented at the *Midwest Conference on Teaching Statistics*, Oshkosh, WI, 1997.

[12] L. A. Saldanha, and F. W. Thompson. Students' reasoning about sampling distributions and statistical inference. In R. Speiser & C. Maher (Eds.), *Proceedings of The Twenty-Third Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. 2001,* pp. 449-454. Snowbird, Utah. Columbus, Ohio: ERIC Clearinghouse.

[13] R. delMas, J. Garfield, and B. Chance. A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education. 1999, 7*(3), [Online]. Available: http://www.amstat.org/publications/jse/secure/v7n3/delmas.cfm]

[14] A. Rubin, B. Bruce, and Y. Tenney. Learning about Sampling: trouble at the Core of Statistics. Paper presented at the *Third International Conference on Teaching Statistics*, New Zealand, 1990.

[15] Weisstein, E. (2003). [Online]. Available::http://www.mathworld.wolfram.com

[16] D. Downing. *Dictionary of Mathematical Terms (2nd ed.).*Hauppauge, NY: Barron's Educational Series, Inc., 1995.

[17] R. Larson, and B. Farber. *Elementary Statistics: Picturing World (2nd ed.).* Upper Saddle River, NJ: Prentice Hall, 2003.

[18] C.A. Barkley. (1995). A study of misconceptions in statistics and statistics anxiety: Affect and performance in an undergraduate math course (Doctoral dissertation, University of Denver). *Dissertation Abstracts International, 56,* AAI9543983.

[19] J.O. Bennett, and. W.L. Briggs. *Using and understanding mathematics: A quantitative reasoning approach (2nd ed.)* New York: Addison Wesley, 2002.

[20] I. Gal. Adults' Statistical Literacy: Meanings, Components, and Responsibilities. *International Statistical Review. 2002, 70* (1).

[21] I. Gal. Statistical Literacy - Meanings, Components, Responsibilities. *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, edited by Ben-Zvi, D. and Garfield, J., Kluwer Academic Publishers. 2004, pp.47-78.

[22] J. Garfield. Thinking about Statistical Reasoning, Thinking, and Literacy. Paper presented at *First Annual Roundtable on Statistical Thinking, Reasoning, and Literacy, 1999.*

[23] J. Garfield. (2002) The Challenge of Developing Statistical Reasoning. *Journal of Statistics Education.* [Online]. Available: www.amstat.org/publications/jse/v10n3/garfield.html

[24] J. Hiebert and P. Lefevre. (1986). Conceptual and Procedural *Knowledge in Mathematics: An Introductory Analysis. in Conceptual and Procedural Knowledge: The Case of Mathematics*, James Hiebert, (ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 1 – 27.

[25] D. Ben-Zvi, and J. Garfield. *Goals, Definitions, And Challenges. The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, edited by Ben-Zvi, D. and Garfield, J., Kluwer Academic Publishers. 2004, pp.3-15

[26] D.J. Rumsey. Statistical Literacy as a Goal for Introductory Statistics Courses. *Journal of Statistics Education*. 2002, Volume 10, Number 3

[27] M. Aliaga, G.W. Cobb, C. Cuff, J. Garfield, R. Gould, R. Lock, et al. (2005). Guidelines for assessment and instruction in statistics education (GAISE) project college report. [Online]. Available: http://it.stlawu.edu/~rlock/gaise/GAISECollege.pdf

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

[29] L. Bude. Assessing students' understanding of statistics. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute, 2006.

[30] T. G. Bond, and C.M. Fox. Applying The Rasch Model: Fundamental Measurement In The Human Sciences. London: Lawrence Erlbaum Associates, 2001.
[31] J. M. Linacre. *A User's Guide To Winstep, Ministep: Rasch ModelComputer Programs.* Chicago: Mesa, 2003.
[32] J.C. Alderson, C. Clapham, and D. Wall. *Language Test Construction And Evaluation*. Cambridge: Cambridge University Press, 1995.
[33] T. Mcnamara. (1996) *Measuring Second Language Performance*. London: Longman. Wong, Harriet; Hazita Azman & Lee Siew Chin. (1990) *English Language Proficiency, Monograph 3*. Kuala Lumpur: Language Centre, Universiti Kebangsaan Malaysia.

IACSIT
International Association of
Computer Science and Information Technology
WWW.IACSIT.ORG