The Construction and Validity of Scale of Self-Efficacy for Statistical Literacy

Su-Wei Lin[1]   Chien-Yi Huang [1]

[1] Department of Education, National University of Tainan, Taiwan

Abstract

This study aims to investigate the development and application of the scale of self-efficacy for senior high school students' statistical literacy (SSESL). Statistical literacy refers to the ability to interpret and critically evaluate messages that contain statistical elements. Regarding science and technology have been playing important roles in our lives; learning probability and statistics is an important topic of modern mathematics education. There were 580 students participated as the norm group. Rasch model was adopted to provide the psychometric characteristics of the SSESL.

Keywords: self-efficacy, statistical literacy, scale of self-efficacy for statistical literacy

1. Introduction

Quantitative information is everywhere, and statistics are increasingly presented as a way to add credibility to advertisements, arguments, or advice. Being able to properly evaluate evidence (data) and claims based on data is an important skill that all students should learn as part of their educational programs. Given the importance of statistics, that there has been an increase in the amount of statistical content included in the elementary and secondary mathematics curriculum and an ever-increasing number of introductory statistics courses taught at the college level (NCTM, 2000). The ability to interpret and critically evaluate messages that contain statistical elements, termed statistical literacy (Gal, 2003), is paramount in our information rich society.

While researchers have investigated student cognitive development in statistical literacy, few have explicitly investigated the associated affective development and its dimensions (Callingham & Watson, 2005; Watson, 2006). Silvia (2003) argued that students with low levels of self-efficacy for a given task are certain that they cannot master it and as a result will lose interest. On the other hand, students with high levels of self-efficacy are certain that they can master the task and as a result will also lose interest. Students with mid-levels of self-efficacy, those with sufficient uncertainty regarding their mastery of the task, are expected to report the most interest.

In this study, students' beliefs regarding their competency are assessed through a construct termed self-efficacy, which is defined as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997). The aim of this study is to develop a self-efficacy scale that will be used in a senior high school statistical literacy context, termed the scale of self-efficacy for statistical literacy (SSESL).

2. Methodology

2.1. Development of Study Tools

In this study, it is expected that students' self-efficacy towards statistical lit-

eracy should reflect the differences of students' statistical illiteracies. Contexts, however, are chosen by teachers to suit the specific needs of their students and can vary widely. As a result, this study has focused on more general contexts, in particular those that are school-, and media-related. A bank of sixty-two items was developed to reflect the statistical topics and contexts used in the teaching of statistical literacy. And this study used a sample of eight items to assess students' self-efficacy with the common stem "How confident do you feel about having to do the following chance and statistical tasks?". Students were required to gauge how closely they could identify with each self-description using a four-point Likert scale ranging from 1 ("Not at all confident") to 4 ("Very confident").

(a) Using a train timetable to work out how long it would take to get from one place to another.
(b) Calculating the average of the second quiz points of your class.
(c) Calculating the probability of being the sample to participate the Basic Language Competency Test
(d) Understanding graphs presented in Newspapers.
(e) using the other of measures (except average) to compare the performance of the two classes
(f) Interpreting the meaning of the probability of rainfall in the weather forecast
(g) Calculating the probability of hitting the jackpot
(h) Interpreting the inconsistency between the results of market survey and the actual selling

2.2. Participants

A total of 624 senior high students from 4 schools in Taiwan, were invited to participate in the study. The results reported here are based on 580 complete responses, a response rate of 93%. Of this sample 57% were male.

2.3. Analysis of Students' Responses

The factor analysis was undertaken using SPSS software to establish and explore possible dimensions within which related items might be located. It employed principal components analysis followed by varimax rotation and estimation of the internal consistency. Throughout the refinement process only those items that loaded on to a single factor with loadings above 0.4 were retained. Such analysis quickly led us to conclude that our decision to assign only tentatively items to dimensions prior to field testing and to let the data and its analysis help to determine the final dimensions had been appropriate.

Student responses were analyzed using the Rasch Model (Andrich, 1978). In addition to the model fit statistics described above, this program also provides an estimate for Separation Reliability. This study reports the infit statistic (denoted MNSQ), which is an information (or variance) weighted sum of squared standardized residuals.

3. Results

Based on extensive review of the literature on self-efficacy for statistical literacy, we developed 12 items reflecting that literature originally. We conducted the principal factor analysis, the result of factor analysis shows that there are one factor loadings of the items identified by their number on the original 8-item version of the instrument and also identified according to their coding on the final version of the SSESL.

Using the Rasch Model of Item Response Theory (IRT) to analyze the students' responses, results showed that the instrument provides a measure that ex-

plains 66% of the variance in student responses with a reported reliability coefficient of α=0.88. Estimates of item difficulties are shown in Table 1. The difficulty estimates for the measure, range from -2.25 to 0.69 logits. The items (f), (g) and (h) are more difficulty than other items. And items (a), (b) and (d) are the simplest. This is not surprising in the case of items (a) and (d) are all about graph, the item (b) is the average problem. The (f), (g) and (h) need higher order critical thinking, it's more difficult than the others for senior high school students'.

Table 1. Rasch rating scale analysis of SSESL

item	Item difficulty	Error	OUTFIT		INFIT	
			MNSQ	ZSTD	MNSQ	ZSTD
(a)	-1.90	0.07	1.05	0.9	1.05	0.9
(b)	-2.25	0.08	0.79	-3.8	0.91	-1.4
(c)	-1.10	0.07	0.88	-2.2	0.89	-2.0
(d)	-1.67	0.08	0.96	-0.7	0.97	-0.4
(e)	-0.94	0.07	0.98	-0.3	0.99	-0.2
(f)	-0.77	0.07	1.09	1.5	1.09	1.5
(g)	0.07	0.06	1.15	2.4	1.17	2.8
(h)	0.01	0.07	1.04	0.8	1.05	0.9

Table 2 presents the summary of 8 Measured Items and overall model fit information. For rating scales, such as the one used in this study, Bond and Fox (2007) recommended that $0.6 \leq MNSQ \leq 1.4$. Based on such criterion, the range of MNSQ is 0.89 to 1.15. It shows that SSESL is fitted Rasch Model.

Table 2. Summary of 12 Measured Items and overall model fit information

Item difficulty M (SD)	Examinees' score M(SD)	Examinees' ability M(SD)	MNSQ	Separation Reliability
-1.07(0.85)	15.26(3.15)	-0.01(0.6)	0.89 ~1.17	0.993

In this study, the correlation between Examinees' ability of self-efficacy and statistical literacy is 0.58. Consequential evidence concerns the future impact that any proposed instrument may have on students who complete the instrument. In this instance it is important that items do not differentiate between sub-groups of students. As reported, item calibrations for males and females were not significantly different at the 1% level. And it's not significantly for examinees' age, too.

4. Discussion

The SSESL was developed following an extensive review of the literature on self-efficacy for statistical literacy and the incorporation of items reflecting that literature into an initial 12-item instrument. This initial instrument was reduced to 8 items following an iterative process of factor analysis and Rasch analysis. The SSESL is an instrument developed for exploring students' self-perceptions of elements of their self-efficacy. We found that items (f), (g) and (h) are more difficulty than other items. And items (a), (b) and (d) are the simplest. This is not surprising in the case of items (a) and (d) are all about graph, the item (b) is the average problem. The (f), (g) and (h) need higher order critical thinking, it's more difficult than the others for senior high school students'.

This study provides one of the examples of the combined use of use of factor and Rasch analysis in the development of such a psychometric instrument in statistics education. The items that constitute them reflect this analysis and are justifiable in terms of their salience in relation what is known from the literature to be important for statistics learning; their usefulness in contributing to understanding students' perceptions of their learning processes; and their confirmatory insights, however minor, in relation to what is already known about the learning context and learning processes of the sample that aided in the development of this instrument.

Finally, we can see the older students had more experience solving problems involving averages (item (b) and (e)); in addition there is some evidence to confirm some research results that suggested boys are more self-efficacious than girls in general (e.g. Pintrich & De Groot, 1990).

5. References

[1] National Council of Teacher of Mathematics. Principals and Standards for School Mathematics. Reston, VA: NCTM, 2000.

[2] Gal, I. "Teaching for statistical literacy and services of statistics agencies." *The American Statistician, 57*(2), pp.80-84, 2003.

[3] Callingham, R., & Watson, J. M. "Measuring Statistical Literacy." *Journal of Applied Measurement, 6*(1), pp.1-29, 2005.

[4] Watson, J. M. Statistical literacy at school: Growth and goals. Mahwah, NJ: Lawrence Erlbaum Associates, 2006.

[5] Silvia, P. J. "Self-efficacy and interest: Experimental studies of optimal incompetence." *Journal of Vocational Behavior, 62*(4), pp.237-249, 2003.

[6] Bandura, A. Self-efficacy: The Exercise of Control. New York: W. H. Freeman, 1997.

[7] Andrich, D. "A rating formulation for ordered response categories." *Psychometrika*, *43*, pp.357-74, 1978.

[8] Bond, T. G., & Fox, C. M. Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum, 2007.

[9] Pintrich, P. R., & De Groot, E. V. "Motivational and self-regulated learning components of classroom academic performance" *Journal of Educational Psychology, 82*(1), pp. 33-40, 1990